

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2003, 05:00:47 ; Search time 4276 Seconds
(without alignments)
3239.695 Million cell updates/sec

Title: US-09-960-643-2
Perfect score: 2513
Sequence: 1 MKRKEDDCSSKKQTTNIR.....VKAGSSSHCRAGQTGVCLIM 476

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09960643/runat_07032003_090517_19877/app_query.fasta_1.647
-DB=GenEmbl -QFMT=fastp -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09960643 @CNC_1_1_2496 @runat_07032003_090517_19877 -NCPU=6 -ICPU=3
-NO_XLPAY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl :
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	2513	100.0	1738	9 HS272L161	AL049688 Human gen
2	2513	100.0	2447	6 AX399682	AX399682 Sequence
3	2513	100.0	2474	9 BC032787	BC032787 Homo sapi
4	2509	99.8	2464	9 AF428261	AF428261 Homo sapi
5	2506	99.7	2612	9 AK095713	AK095713 Homo sapi
6	2327.5	92.6	2416	10 BC021840	BC021840 Mus muscu
7	2327.5	92.6	2427	10 AF428262	AF428262 Mus muscu
8	1612	64.1	1013	10 D86557	D86557 Rattus norv
9	1246	49.6	1074	6 AX166517	AX166517 Sequence
10	1246	49.6	1074	6 AX167587	AX167587 Sequence
11	1246	49.6	1158	6 AX167585	AX167585 Sequence
12	1246	49.6	1579	9 AF286366	AF286366 Homo sapi
13	1246	49.6	1671	6 AX167589	AX167589 Sequence
14	1157	46.0	1448	10 BC014825	BC014825 Mus muscu
15	1155.5	46.0	1439	10 RATCAMPKAA	L22888 Rattus norv
16	1154	45.9	1480	9 HUMCKI	L41816 Homo sapien
17	1151	45.8	1402	10 RATPRKI	L24907 Rattus norv
18	1070	42.6	1618	10 AB004267	AB004267 Rattus no
19	1068	42.5	1332	10 D86556	D86556 Rattus norv
20	1065	42.4	1554	10 AF181984	AF181984 Mus muscu
21	1064	42.3	1032	10 AB023027	AB023027 Mus muscu
22	1041	41.4	1032	6 AX166520	AX166520 Sequence
23	1004	40.0	1546	3 AB021864	AB021864 Caenorhab
24	990	39.4	1282	6 AR139102	AR139102 Sequence
25	957.5	38.1	3316	3 DMEI7917	Y17917 Drosophila
26	865	34.4	157875	9 HS272L16	AL023754 Human DNA
27	840	33.4	2202	9 BC033746	BC033746 Homo sapi
28	809.5	32.2	42404	3 CBRC46106	AC084656 Caenorhab
29	753	30.0	2897	6 AX099935	AX099935 Sequence
30	747.5	29.7	1503	6 AX239863	AX239863 Sequence
31	747.5	29.7	2297	6 AX239861	AX239861 Sequence
32	747.5	29.7	3003	9 BC001921	BC001921 Homo sapi
33	747.5	29.7	3012	9 BC000497	BC000497 Homo sapi
34	747.5	29.7	3012	9 BC005828	BC005828 Homo sapi
35	747.5	29.7	3018	9 BC017363	BC017363 Homo sapi
36	747.5	29.7	3057	9 BC019256	BC019256 Homo sapi
37	747.5	29.7	3124	6 AX350342	AX350342 Sequence
38	746	29.7	3141	10 BC017634	BC017634 Mus muscu
39	741	29.5	3655	10 RATCBVA	L22557 Rattus norv
40	735	29.2	182054	10 AL365314	AL365314 Mouse DNA
41	735	29.2	184754	2 AC022675	AC022675 Mus muscu
42	699.5	27.8	955	3 DDIMLCK	M64176 D.discoideu
43	697.5	27.8	3165	10 MMCDPK4	Y58995 Mouse mRNA
44	692.5	27.6	1706	10 S65840	S65840 Ca2+/calmod
45	692.5	27.6	1785	10 RATCALM	M64757 Rat Ca++/ca

ALIGNMENTS

RESULT 1

```

HS272L161
LOCUS       HS272L161              1738 bp    mRNA    linear    PRI 21-APR-1999
DEFINITION   Human gene isolated from PAC 272L16, chromosome 1, similar to
              calcium/calmodulin dependent protein kinases.
ACCESSION    AL049688
VERSION      AL049688.1  GI:4678721
KEYWORDS     .
SOURCE       Homo sapiens.
ORGANISM     Homo sapiens.
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
TITLE        1 (bases 1 to 1738)
JOURNAL      Rhodes, S.
COMMENT      Direct Submission
              Submitted (21-APR-1999) E-mail contact: humquery@sanger.ac.uk
              This sequence was generated from cDNA clones isolated using
              sequence from the bacterial clone 272L16 (AL023754) and EST data.
              The EST sequences listed match this sequence with an identity of at
              least 95% between the coordinates shown.
              Further information can be found at
              http://www.sanger.ac.uk/HGP/Chr1/ Partial, experimentally
              determined gene.
FEATURES             Location/Qualifiers
     source          1..1738
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="lq32.1-32.3"
     CDS             1..1448
                     /codon_start=3
                     /product="hypothetical protein"
                     /protein_id="CAB41259.1"
                     /db_xref="GI:4678722"
                     /translation="STLEAMGRKEEDDCSSWKQTTNIRKTFIFMEVLGSGAFSEVFL
                     VKQRTGLKFAKLCIKKSPAPRDSLENEIAVLKKIKHENIVTLEDIYESTHYLVLM
                     QLVSGELDRILRGVYTEKDALSVIQOVLAVKYLHENGIVHRDLKAPENLLLTPE
                     ENSKIMITDFGLSKMEQNGIMSTAGTPEGVAPVLAQPKYSKVDGWSIGVIYIILL
                     CGYPPEYESKLEFKIEGYEFESFPWDDISEAKDFICHLEKDPNERYTCEKA
                     LSHPTWNTALHRDIYPSVLIQKNFAKSKWQAFNAAVYVHHMKLHMLNLSPGV
                     RPEVENRPPQASSTSRPSPSEITITEAPVLDSHVALPALPOLPCQHRPRPTAPGGR
                     SLNCLVNGSLHTSSSLVPMHOGSLAAGPCGSCSLNIGSKGKSSYCSSEPTLLKKANK
                     KQNFSEVMVPVKASGSHCRAGQVGLIM"
                     <1..107
                     /number=1
                     /number=2
                     /number=3
                     /number=4
                     /number=5
                     /number=6
                     /number=7
                     /number=8
                     /number=9
                     /number=10
                     /number=11
                     /number=12
     misc_feature    653..777
                     /note="matches EST AI215131 from clone IMAGE:1925595"
     exon            764..842
                     /number=1
     exon            843..930
                     /number=2
     exon            868..1159
                     /number=3
     exon            931..1355
                     /number=4
     exon            1356..1483
                     /number=5
     misc_feature    join(1469..1550,1553..1723,1721..1738)
                     /note="matches EST R05661 from clone 29500"
     exon            1484..1736
                     /number=6
     exon            1484..1736
                     /number=7

```



```
Qy 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeu 320
|
|
|
Db 918 AAGAGCAAGTGGAGGCAAGCTTCAACGAGCAGAGTGTGGTGCACCACATGAGGAAGCTA 977
|
|
|
Qy 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProGluThr 340
|
|
|
Db 978 CACATGAACCTGGACAGCCGGCGCTCCGCCAGAGGTGGAGAACAGCGCGCTGAAACT 1037
|
|
|
Qy 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360
|
|
|
Db 1038 CAACCTCAGAACCTCTAGACCAGCTCCCTGGAGATCACCATCACCGAGGACCTGTC 1097
|
|
|
Qy 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
|
|
|
Db 1098 CTGACACACAGTGTAGCACCTCCCTGGCTGACCAATTACCTGCCAGCATGGCGCGCG 1157
|
|
|
Qy 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
|
|
|
Db 1158 CCCACTGCCCTGGTGCAGGTCCCTCACTGGCTGGTCAATGGCTCCCTCCACATCAGC 1217
|
|
|
Qy 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420
|
|
|
Db 1218 AGCAGCTGGTGGCCATGATCAGGGTCCCTGGCGCGCGGCGCTGGCTGCCCTC 1277
|
|
|
Qy 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerSerSerSerSerSerSerSerSer 440
|
|
|
Db 1278 AGCTGCTGAACATGGAGCAAGGAAGTCTCTCTACTGCTGCTGAGCCCACTCCCTC 1337
|
|
|
Qy 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460
|
|
|
Db 1338 AAAAGGCCAACAAAAACAGACTCAAGTCGAGGTCATGGTACCAGTTAAAGCCAGT 1397
|
|
|
Qy 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
|
|
|
Db 1398 GGCAGCTCCCACTGCCGGCAGGCGACACTGGAGTGTGCTCATATTATG 1445
|
|
|
RESULT 2
AX399682 2447 bp DNA linear PAT 06-JUN-2002
LOCUS
DEFINITION Sequence 3 from Patent WO0224947.
ACCESSION AX399682
VERSION AX399682.1 GI:21335455
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
1 Delaney,A.D. and Yoganathan,T.
Cancer associated protein kinases and their uses
Patent: WO 0224947-A 3 28-MAR-2002;
JOURNAL KINETEK PHARMACEUTICALS INC (CA); UNIV BRITISH COLUMBIA (CA)
FEATURES
source
1. .2447
/organism="Homo sapiens"
/db_xref="taxon:9606"
70. .1500
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD33521.1"
/db_xref="GI:21335456"
/translation="MGRKEEDDCSSKKQTTNIRKTFIMFVLGSGAFSEVFLVKQRL
TGKLFALKTKKSPAFRDSLENIATLKKIHENIVLIEDIYESTTHYIVLMQLVSG
GELFDRLRGVYTKEDASIVTQQVL SAVKY LHENGIVHRDLKPNLLYLPENSKI
MITDFGLSKMEONGIMSTAGCTPGVVAPEVLAQKPYSKAVDCWSITGVITYILCGYPP
FYETESKLPFKIEGYVEFSPFWDLLSESAKDFICHLLEKDPNERYTCEKALSHPW
IDNTALHRDIDPVSLSIQKNFAKSKWRQAFNAAYVHHMKLHMNLHSGRPREVE
NRPEFOAETSRPSPEITTEAPVLDHVSVALPALQTLPQOHRHRTAPGRSVNLCI
VNGLSHLSLSPVHMQGSIAAGPCGCCSLNIGSKGSSYCSSEPTLLKKANKKQNEK
SEYVNPVKAGSGSHCRAGGTGVCILIN"
BASE COUNT 590 a 707 c 604 g 546 t
ORIGIN
```

```
Alignment Scores: 3.62e-170 Length: 2447
Pred. No.: 2513.00 Matches: 476
Score: 2513.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 6 Gaps: 0
DB:
US-09-960-643-2 (1-476) x AX399682 (1-2447)
Qy 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
|
|
|
Db 70 ATGGGTGAAAGGAAGAGATGACTGCAGTTCCTGGAGAAACAGACCACCAATCCCG 129
|
|
|
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
|
|
|
Db 130 AAAACCTTCATTTTATGAGTCTGGGATCAGGAGCTTCTCAGAAAGTTTTCCTGGTG 189
|
|
|
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
|
|
|
Db 190 AAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAGTCACCTGCTTC 249
|
|
|
Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
|
|
|
Db 250 CGGACACAGCCTGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAACATTT 309
|
|
|
Qy 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
|
|
|
Db 310 GTGACCTGGAGGACATCTATGAGACCAACCACCTACTACTGCTGTCATGCAGCTTGT 369
|
|
|
Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
|
|
|
Db 370 TCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAGAGATGCC 429
|
|
|
Qy 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
|
|
|
Db 430 AGTCTGTGTATCCAGCAGCTTGTGCGCAGTGAATACCTACATGAGAAATGCGCATCGTC 489
|
|
|
Qy 141 HisArgAspLeuLysProGluAsnLeuTyrLeuThrProGluGluAsnSerLysIle 160
|
|
|
Db 490 CACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTTACCCCTGAAGAACTCTTAGATC 549
|
|
|
Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
|
|
|
Db 550 ATGATCATCTGACTTTGCTGCTCCACAGATGAACAGAAATGCGCATCATGCTCCTGCTGT 609
|
|
|
Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
|
|
|
Db 610 GGGACCCCGAGCTACGTGGCTCCAGAAAGTGTGGCCCGAGAAACCCCTACAGAGGCTGTG 669
|
|
|
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220
|
|
|
Db 670 GATTGCTGGTCCATCGCGCTCATCCTACATATGCTCTGTGGATACCCCGCTTCTAT 729
|
|
|
Qy 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
|
|
|
Db 730 GAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGATGTTTGTGCTCT 789
|
|
|
Qy 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
|
|
|
Db 790 CCAATCTGGGATGACATTTCTGAGTCAGCCCAAGGACCTTATTTGCCACTTCTTGAGAAG 849
|
|
|
Qy 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
|
|
|
Db 850 GATCCGAACGAGCGGTACACCTGTGAGAAGCCCTTGAGTCATCCCTGCATGATGACGAAAC 909
|
|
|
Qy 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
|
|
|
Db 910 ACGGCCCTCCACCGGGACATCTACCCATCAGCTCAGCTCCAGATCCAGAGAACTTTGCT 969
|
|
|
Qy 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeu 320
|
|
|
Db 970 AAGACCAAGTGGAGGCAAGCCTTCAACGACAGCAGCTGTGGTGCACCACATGAGGAAGCTA 1029
|
|
|
```

```
Qy 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr 340
|||||
Db 1030 CACATGAACCTGCACACCGCGGCTCGGCCAGAGGTGGAGAACAGCGCGCTGAACCT 1089
|||||
Qy 341 GlnAlaSerGluThrSerArgProSerSerProGluLeuThrIleThrGluAlaProVal 360
|||||
Db 1090 CAAGCCCTCAGAAACCTTAGACCCAGCTCCCTCGATCACCATCACCAGGACCATGTC 1149
|||||
Qy 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
|||||
Db 1150 CTGACCACAGTAGACATCCCTGCGCTGACCAATACCTGCGCAGATGGCGCGCG 1209
|||||
Qy 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
|||||
Db 1210 CCCACTGCCCTGGTGCAGGTCCTCAACTGCCTGCTCAATGGCTCCCTCCACATCAGC 1269
|||||
Qy 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420
|||||
Db 1270 AGCAGCTGTGTCATGATCATCAGGGTCCCTGGCGCGCGGCTGTGCTGCTGCTCC 1329
|||||
Qy 421 SerCysLeuAsnIleGlySerGlyLysSerSerTyrCysSerGluProThrLeuLeu 440
|||||
Db 1330 AGTGGCTGAACATTGGGACCAAGAAAGTCTCTACTGCTCTGAGCCACACTCTCTC 1389
|||||
Qy 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460
|||||
Db 1390 AAAAAAGCCCAACAAAAACAGAACTTCAAGTCGGAGGTGATGTATCAGTTAAAGCCAGT 1449
|||||
Qy 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
|||||
Db 1450 GGCAGTCCCACTCGCGGCGAGGACAGTGGAGTGTGCTCATATTG 1497
|||||
RESULT 3
BC032787 2474 bp mRNA linear PRI 27-JUN-2002
LOCUS Homo sapiens, calcium/calmodulin-dependent protein kinase IG, clone
DEFINITION MGC:44894 IMAGE:5179957, mRNA, complete cds.
ACCESSION BC032787
VERSION BC032787.1 GI:21619664
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2474)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Center by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nih.gov
Akhtar, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
```

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 68 Row: k Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14196444.

```
FEATURES
    source
    location/Qualifiers
    1..2474
    /organism="Homo sapiens"
    /db_xref="LocusID:57172"
    /db_xref="taxon:9606"
    /clone="MGC:44894 IMAGE:5179957"
    /tissue_type="Brain, Lung, Testis, adult, pooled whole"
    /clone_lib="NIH_MGC_115"
    /lab_host="DH10B"
    /note="Vector: pCMV-SPORT6"
    70..1500
    /codon_start=1
    /product="calcium/calmodulin-dependent protein kinase IG"
    /protein_id="AAH32787.1"
    /db_xref="GI:21619665"
    /translation="MGRKEEDCSWKKQTNTIRKTFIFMEVLGSGAFSEVFLVKORL
    TGLFALKCIKKSPAFRSDSLENEIAVLKKIKHENIVLHDLKPENLLYLTPENSKI
    GELFDRLERGVYTEDASLVIOQVL SAVKYHENGIVHRLDKPENLLYLTPENSKI
    MINDFGLSKMEQNGIMSTAGTGYVAPEVLAOKPYSKAVDCWSIGVITVILCGYPP
    FYEETESKLEFKIEGYEYESPFWDIDISESAKDFICHLEKDPNERVTCERKALSHPM
    IDGNTALURDIIYFSVLQIONFAKSKWRQFNAAAVVHHMKLHNNLHNSPGVRPEVE
    NRPPETQASETSPPSEITITEAPVLDHSEVALPALTQLPCQHGRRPTAPGGHSLNCL
    VNGSLHSSSLVPMHQGSLAAGCGCCSCLNIGSKGKSSCYCSEPTLLKANKKKQNFK
    SEVMVPVKASGSSHCRAGQTGVCILM"
    BASE COUNT 618 a 707 c 603 g 546 t
    ORIGIN
    Alignment Scores:
    Pred. No.: 3.67e-170 Length: 2474
    Score: 2513.00 Matches: 476
    Percent Similarity: 100.00% Conservative: 0
    Best Local Similarity: 100.00% Mismatches: 0
    Query Match: 100.00% Indels: 0
    DB: 9 Gaps: 0
    US-09-960-643-2 (1-476) x BC032787 (1-2474)
    Qy 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
    Db 70 ATGGGTCCGAAGGAAGAGTGCATGCTGCTCGGAAGAAGACAGACCACCAATCCGG 129
    Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
    Db 130 AAAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTTCTCAGAAGTTTCTCGTG 189
    Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
    Db 190 AAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCAATCAAGAAGTCACTGCCTTC 249
    Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
    Db 250 CGGACACAGCTGGAGAAATGAGATGCTGTTGAAAAGATCAAGCATGAACATT 309
    Qy 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
    Db 310 GTGACCTTGGAGGACATCTATGAGAGCACCACCTACTACTCTGTCATGCGCTGTT 369
    Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
    Db 370 TCTGTTGGGAGCTCTTTGACCGGATCTCGAGCGGGGTGCTACACAGAGAAGGATGCC 429
    Qy 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
    Db 430 AGTCTGGTATCCAGCAGCTCTTGTGCGCAGTGAATACCTACATGAGAATGGCATCGTC 489
    Qy 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
    Db 490 CACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCTTACCTTACCTTACCTTACCTT 549
```

```
Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 550 ATGATCACTAGCTTTGGTCTGTCCAAGATGGAACAGAAATGCCATCATGTGCCACTGCTGT 609
Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 610 GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGCCAGAAACCCCTACAGCAAGGCTGTG 669
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220
Db 670 GATTGCTGGTCCATCGCGCTCATCACCTACATATTTGCTGTGTGATACCCCGGTCTAT 729
Qy 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Db 730 GAAAGAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTTACTATGAGTTGAGTCT 789
Qy 241 ProPheTrpAspPheLeuSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
Db 790 CCATCTGGGATGACATTTCTGAGTCAGCCAGGACTTTATTGGCCACTTGTCTGAGAG 849
Qy 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
Db 850 GATCCGACGAGCGGTACACTGTGAGAGGCTTGTGAGTCCATCCCTGGATTGACGGAAC 909
Qy 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
Db 910 ACAGCCTCCACCGGGACATCTACCCATCATGCTCAGCTCCAGATCCAGAGAACTTTGCT 969
Qy 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeu 320
Db 970 AAGAGCAAGTGGAGGCAAGCTTCAAGCGCAGCAGCTGTGGTGCACCATGAGGAGCTA 1029
Qy 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr 340
Db 1030 CACATGAACTGCGACAGCCCGGGCTCCGCCAGAGGTGGAGAACAGGCGCGCTGAACT 1089
Qy 341 GlnAlaSerGluThrSerArgProSerProGluIleThrIleThrGluAlaProVal 360
Db 1090 CAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACCTGTC 1149
Qy 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
Db 1150 CTGACACACAGTGTAGCAGCTCCCTGGCCCTGACCCAAATTACCCTGCAGCATGCCGCCG 1209
Qy 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
Db 1210 CCCACTGCCCTGTGGCAGGTCCCTCACTGCTGGTCAATGCTCCCTCCACATCAGC 1269
Qy 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420
Db 1270 AGCAGCTGTGTGCCATGCATCAGGGTCCCTGCGCGCGCGGCTGTGCTGCTGCTCC 1329
Qy 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu 440
Db 1330 AGCTGCTGACATTTGGAGCAAGGAAGTCTCTCTACTGCTCTGAGCCCACTCCCTC 1389
Qy 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460
Db 1390 AAAAAGGCCAACAAAAACAGACTTCAAGTCGAGGTGATGTTACAGTTAAAGCCACT 1449
Qy 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
Db 1450 GGCAGCTCCCACTCCGCGGAGGCGAGCTGGAGTGTGTCTCATATG 1497

RESULT 4
AF428261
LOCUS AF428261 2464 bp mRNA linear PRI 07-NOV-2001
DEFINITION Homo sapiens calcium/calmodulin-dependent protein kinase I gamma
(CAMK1G) mRNA, complete cds.
ACCESSION AF428261
VERSION AF428261.1 GI:16755791
KEYWORDS
SOURCE Homo sapiens.
```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2464)
Schutte,B.C., Bjork,B.C., Coppage,K.B., Malik,M.I., Gregory,S.G.,
Scott,D.J., Brentzell,I.M., Watanabe,Y., Dixon,M.J. and Murray,J.C.
A preliminary gene map for the Van der Woude syndrome critical
region derived from 900 kb of genomic sequence at 1q32-q41
Genome Res. 10 (1), 81-94 (2000)
20113118
MEDLINE
PUBMED 10645953
REFERENCE 2 (bases 1 to 2464)
Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.
Characterization of the human ortholog of rat Cam Kinase I gamma
(CamK1g) at 1q32-q41
AUTHORS Unpublished
JOURNAL 3 (bases 1 to 2464)
Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.
Direct Submission
AUTHORS
TITLE
JOURNAL
SUBMITTED (05-Oct-2001) Genetics, University of Iowa, 140 EMBR,
Iowa City, IA 52242, USA
FEATURES
Location/Qualifiers
1. .2464
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1q32-q41"
1. .2464
/gene="CAMK1G"
73. .1503
/gene="CAMK1G"
/codon_start=1
/product="calcium/calmodulin-dependent protein kinase I
gamma"
/protein_id="AAL28100.1"
/db_xref="GI:16755792"
TGKLFALKICKKSPAPRDSLEIAVLKKIKHENITVLTIEDIVESTHYLYVMQLVSG
GELFDRLRGVPTKDSALVIOQLSVAKYKLVHENGIVHDLKPNLLVLTPEENSKI
MITDGLSKMEONGINSTAGCTGGYVAPVLAQPKYSKAVDCWSIGVITYIILCGYPP
FYETESKLFKEIKEGYEPFDDDISAKDFICHLEKDPNRYTCERALSHPW
IDENTALHRDIYPSVLSQKFNKFAKSKRQAFNAAVVHMRKLNHMLHSPGIRPEVE
NRPETQASETSREPSPEITTEAPYLDHSVALPALTQPCQHRRTAPGGRSLNCL
VNGSLHISSSLVPMHOGSLAAGPCGCCSCLNIGNKGSKSYCSEPTLLKANKKQNFK
SEVMVPVKASGSSSHCRAGGTGVCILM"
2415. .2420
polyA_signal
BASE COUNT 609 a 708 c 602 g 545 t
ORIGIN
Alignment Scores:
Pred. No.: 7,05e-170 Length: 2464
Score: 2509.00 Matches: 474
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.58% Mismatches: 0
Query Match: 99.84% Indels: 0
DB: 9 Gaps: 0
US-09-960-643-2 (1-476) x AF428261 (1-2464)
Qy 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrAsnIleArg 20
Db 73 ATGGGTCGAAAGGAAGAAGATGACTGCGAGTTCTCTGGAGAAACAGACCAACATCCG 132
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 133 AAAACCTTCATTTTATGGAAGTGTGGATCAGGAGCTTTCTCAGAAGTTTCTCCTGGTG 192
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 193 AAGCAAGACACTGACTGGGAGACCTCTTGTCTCTGAAGTCATCAAGAGTACCTGCCTTC 252
Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80

```
Db 253 CGGACACGACCGCTGGAGAAATGAGATTGCTGTGTGAAAAAGATCAAGCATGAAACATT 312
Qy 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 313 GTGACCTGGAGGACATCTATGAGAGACACACCCATCTACTACCTGGTGTGTCGACCTGGT 372
Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Db 373 TCTGGTGGGAGCTTTTGACCGGATCTCGAGCGGGTGTCTACACAGAGAAGGATGCC 432
Qy 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 433 AGTCTGTGTATCCAGCAGGCTTGTTCGGCAGTGAATATCTATACATGAGAATGTCATCGTC 492
Qy 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 493 CACAGAGACTTAAAGCCGAAACCTGCTTACCTTACCCCTGAAGAGAACTCTAAGATC 552
Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaLys 180
Db 553 ATGATCACTGACTTGTGCTGTCCAGATGGAACAGATGGCATGTCATGCTCCACTGCCTGT 612
Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 613 GGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCAGAAACCCCTACAGCAAGGCTGTG 672
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLysGlyTyrProProPheTyr 220
Db 673 GATTGCTGGTCCATCGGCGTCATCCCTACATATGCTGCTGGATACCCCGCTCTAT 732
Qy 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Db 733 GAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGTTTGAGTCT 792
Qy 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
Db 793 CCATTCTGGGATGACATTTCTGAGTCAGCAAGGACTTTATTGGCCACTGCTTGAGAAG 852
Qy 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
Db 853 GATCCGAACGAGCGGTACACTGTGAGAAGCGCTTGAGTCATCCCTGGATTGACGGAAC 912
Qy 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
Db 913 ACAGCCCTCCACCGGGACATCTACCCATCAGTCAGCCTCCAGATCCAGAGAATTTGCT 972
Qy 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320
Db 973 AAGAGCAAGTGGAGGCAAGCCTTCAACGCAGCAGCGCGTGTGCACCATGAGGAAGCTA 1032
Qy 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProGluThr 340
Db 1033 CACATGAACCTGCACAGCGCGGATCCGCCAGAGTGAGAGTGGAGAACGCGCTGAAACT 1092
Qy 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360
Db 1093 CAACCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCACCGAGGCACTGTC 1152
Qy 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
Db 1153 CTGGACACAGTGTAGCACTCCCTGCCCTGACCAATTAACCTGCCAGCATGCGCCGCG 1212
Qy 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
Db 1213 CCCACTGCCCTGTGTGCGAGTCCCTCAACTGCTGTGCTCAATGGCTCCCTCCACATCAGC 1272
Qy 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420
Db 1273 AGCAGCTGTGTGCCATGATCAGGGGTCCCTGGCCCGCGGCGCTGTGGCTGCTGCTCC 1332
Qy 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu 440
```

```
Db 1333 AGCTCCCTGAACATTTGGGAACAAGGAAAGTCTCTCTACTGCTGTGAGCCACACTCCTC 1392
Qy 441 LysLysAlaAsnLysLysGlnAsnPhelySerGluValMetValProValLysAlaSer 460
Db 1393 AAAAAGGCCAACAAAAACAGACTTCAAGTCGGAGGTCAATGGTACCAAGTTAAAGCCAGT 1452
Qy 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
Db 1453 GGCACGCTCCCACTCGCGGCGAGGCGAGACTGGAGTCTGTCTCATTTATG 1500

RESULT 5
AK095713 2612 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ38394 fis, clone FEBRA2007534, moderately
DEFINITION similar to CamKI-like protein kinase.
ACCESSION AK095713 GI:21755035
VERSION AK095713.1 (full insert sequence).
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens fetal brain cDNA to mRNA, clone_lib:FEBRA2
clone:FEBRA2007534.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M.,
Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
REFERENCE 2 (bases 1 to 2612)
AUTHORS Isogai,T. and Yamamoto,J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Source 1. 2612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FEBRA2007534"
/tissue_type="brain"
/clone_lib="FEBRA2"
/dev_stage="fetal"
/note="cloning vector: pME18SFL3"
BASE COUNT 619 a 750 c 645 g 598 t
ORIGIN
Alignment Scores: 1.24e-169 Length: 2612
Pred. No.: 2506.00 Matches: 475
Score: 2506.00 Conservative: 0
Percent Similarity: 99.79% Mismatches: 1
Best Local Similarity: 99.79% Indels: 0
Query Match: 99.72% Gaps: 0
DB: 9
US-09-960-643-2 (1-476) x AK095713 (1-2612)
```

```
Qy 1 MetGlyArgLysGluLysAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 246 ATGGGTCGAAAGGAAGATGACTGCATGCTCTCTGGAAGAAACAGACCACCAATCCGG 305
```


/translation="MGRKEBEDCSSWKKQTNIRKTFIFMEVLGSGNFASEVFLVKORY
TGKFAKCIKSAFADSSLENEIAIKRIKHENIVITLDEIYTHIYLVWQLVSG
GELFDRLERGVYTKDASLVIOQVLSAVKYLHENGIVHRDLAPENLITLTPENSKI
MITDFGLSKMQNGVMSTACGTGYVAPELVAOKPYSAVDWCSIGVITYILLCGYPP
FYETESKLFEKIREGYEFEPFWDIDISEAKDFICHLLEKDPNERYTEKALRHPW
IDGNTALHRDITPVSQIQKNFAKSWROAFNAAVVHMRKLMHNLHSYRQOEVE
NRPVSPAPEVSRDSDSSYTEAPILDPTPLPALPCLPSHSSRPSAPSGRSLNC
LYNSLRITSSSLVPMQOGLATGPGCCSCLNIGPKGSKSYSEPTLFRKANKQNF
KSEWNPVAGGSHCRGGGTGVLVM"
BASE COUNT 572 a 694 c 602 g 548 t
ORIGIN

Alignment Scores:
Pred. No.: 6, 24e-157 Length: 2416
Score: 2327.50 Matches: 439
Percent Similarity: 94.97% Conservative: 14
Best Local Similarity: 92.03% Mismatches: 23
Query Match: 92.62% Indels: 1
DB: 10 Gaps: 1

US-09-960-643-2 (1-476) x BC021840 (1-2416)

QY	1	MetGlyArgGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg	20
Db	87	ATGGGCGTAAAGGAGGAGGACTGCAGTTCTCGAAGAAACAGACCACCAACATCAGG	146
QY	21	LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal	40
Db	147	AAACCTTCATCTTCATGGAAGTGTGGATCAGAGCTTCTCAGAGGTGTTCCTGGTG	206
QY	41	LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe	60
Db	207	AAGCAAGAGTGACTGGGAACCTTTGCTCTGAAATGTATCAAGAACTCACCGCCCTC	266
QY	61	ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle	80
Db	267	CGGACAGCAGCTAGAGAAATGAGATCGCTGTGTTGAAGAGATCAAGCATGAGAACAT	326
QY	81	ValThrLeuGluAspIleTrpGluSerThrHisTrpLysTrpValMetGlnLeuVal	100
Db	327	GTGACCTGGAGGACATCTATCAGAGCACACCCACTACTACCTGGTCATGAGCTTGT	386
QY	101	SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValThrGluLysAspAla	120
Db	387	TCGGAGGTGAGCTTTTGACCGGATCTTAGAGCTGTGTCTACACAGAAAGATGCC	446
QY	121	SerLeuValIleGlnValLeuSerAlaValLysTrpLeuHisGluAsnGlyIleVal	140
Db	447	AGCTGTGTCATCCAGAGCTGTGTCTGGGTGAATACCTTCATGAGAAATGGCATCGTC	506
QY	141	HisArgAspLeuLysProGluAsnLeuLeuThrProGluGluAsnSerLysIle	160
Db	507	CACAGAGATCTAAAGCTGAAACCTGTGTACCTCACCCCTCAGAGAGATCCCAAGATC	566
QY	161	MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys	180
Db	567	ATGATCACCTGACTTTGGTCTATCCAGATGGAGAGATGAGTCTATGTCACAGCTTGT	626
QY	181	GlyThrProGlyTrpValAlaProGluValLeuAlaGlnLysProTrpSerLysAlaVal	200
Db	627	GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAACCCCTACAGTAAGGCTGTG	686
QY	201	AspCysTrpSerIleGlyValIleThrTrpIleLeuLeuLysGlyTrpProPheTrp	220
Db	687	GACTGTGTGTCATGTCATACATACATGCTGTGTGGTATCCCCCTTTCAT	746
QY	221	GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTrpTrpGluPheGluSer	240
Db	747	GAGAAACAGATCAAGCTTTTGGAGAGATCAAGAGGTTTACTACAGTTTTCAGTCT	806
QY	241	ProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys	260
Db	807	CCGTTCTGGGATGACATTTCTGAGTCAGCAAGGATTTTATTTGTCATCTGTCGAGAAG	866

QY	261	AspProAsnGluArgTrpThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn	280
Db	867	GACCCAAACGAGCGGTACCTCGGAAAGCCCTCAGACACCCCTGGATGATGAAC	926
QY	281	ThrAlaLeuHisArgAspIleTrpSerValSerLeuGlnIleGlnLysAsnPheAla	300
Db	927	ACAGCCCTGACCGGACATCTATCCATCTGTGCTCAGCTCCAGATTCAGAAGAACTT	986
QY	301	LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeu	320
Db	987	AAGAGCAAGTGGAGCAAGCCTTCAACGCGAGCTGCGGTGTGTCATCAGAGAACTG	1046
QY	321	HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr	340
Db	1047	CACATGAACCTCAGACGCCCCAGTGTCCGCAAGAGGTGGAGAACAGGACCTGTGTC	1106
QY	341	GlnAlaSerGluThrSerArgProSerProGluIleThrIleThrGluAlaProVal	360
Db	1107	CCTGCTCCTCAAGTCTCCAGACCCAGACTCCCATGACAGCTCCATCAGAGGCCCATC	1166
QY	361	LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg	380
Db	1167	CTGACCCCAAGCACACCCCTTCTGCTGACTGACCCGACTACCTGCTCAGACAGCTCC	1226
QY	381	ProThrAlaPro--GlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIle	399
Db	1227	CCITCAGCTCCAGTGTGTGCGCTCCTCACTGCTGTGTCACAGCGCTCCCTGCGCATC	1286
QY	400	SerSerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysCys	419
Db	1287	AGCAGTAGCTAGTGTGCTGCAACAGGCGCCCTAGCACGGGGCCCTGTGCTGCTGC	1346
QY	420	SerSerCysLeuAsnIleGlySerLysGlyLysSerSerTrpCysSerGluProThrLeu	439
Db	1347	TCAGCTGTCTAAATATTGGGAACAGGGAAGTCTTCTACTGCTGAGCGCTACCTC	1406
QY	440	LeuLysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAla	459
Db	1407	TTCAGAAAGGCCAATAAACAACAGAACTTCAAGTCAGAGGTGATGTATACAGTGAAG	1466
QY	460	SerGlySerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet	476
Db	1467	GTTGGGACACCCACTGCTCCGGGGTGGGACACTGGGGGTGTCTCGTTATG	1517
RESULT 7			
AF428262		2427 bp	mRNA linear ROD 07-NOV-2001
LOCUS		Mus musculus calcium/calmodulin-dependent protein kinase I gamma	
DEFINITION		(Camk1g) mRNA, complete cds.	
ACCESSION		AF428262	
VERSION		AF428262.1	GI:16755793
KEYWORDS			
SOURCE		Mus musculus	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE		1 (bases 1 to 2427)	
JOURNAL		Characterization of the human ortholog of rat Cam Kinase I gamma	
REFERENCE		Unpublished	
AUTHORS		2 (bases 1 to 2427)	
TITLE		Bjork, B.C., Watanabe, Y., Murray, J.C. and Schutte, B.C.	
JOURNAL		Submitted (05-OCT-2001) Genetics, University of Iowa, 140 EMRB,	
FEATURES		Iowa City, IA 52242, USA	
source		Location/Qualifiers	
		1. 2427	
		/organism="Mus musculus"	
		/db_xref="taxon:10090"	
		/chromosome="1"	
		/map="1H, 104.00 cM"	

```
/clone="IMAGE:4502479"
/tissue_type="retina"
/notes="corresponds to EST BG293660"
1..2427
/genes="Camk1g"
86..1519
/genes="Camk1g"
/codon_start=1
/product="calcium/calmodulin-dependent protein kinase I
gamma"
/protein_id="AAL28101.1"
/db_xref="GI:16755794"
/translation="MGRKEEDCSKKKQTTNIRKTFIMEVLGSGARSEVFLYKQRV
TGKLFALKIKKSPAPRDSLENEIAVLKRIKHENIVILEDIESTHYLYVMQLVSG
GEUFLRILRGVTEKDALSLVQQVLSAVKYLVHENGIVHRDLKPNLLYLPPENSKI
MITDFGLSRMEONGVMSACGTPGYVAPVLAOKPYSKAVDCWSIGVATYILCGYPP
FYETESKLFIEKIKEYEYEFSPFWDIDSESADFICHLLEKDPNERTYCEKALRHWP
IDGNTALHRIEIVPSVSLQIQKNFAKWRQAFNAAVVHHMKLHMNHSVSRQVEE
NRPPVSPAPESVRPSDSSITFAPILDPSTPLPALTRLPCSHSRPSAPSGRSLNC
LVNGSIRISSLVPMQGPPLATGPGCCSCNLNIGKSKSSVCSEPTLFRKANKQNF
KSEVMVPVRAGGSTHCRGGQTGVCLVM"
BASE COUNT 584 a 694 c 601 g 547 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 6,27e-157 Length: 2427
Score: 2327.50 Matches: 439
Percent Similarity: 94.97% Conservative: 14
Best Local Similarity: 92.03% Mismatches: 23
Query Match: 92.62% Indels: 1
DB: 10 Gaps: 1

US-09-960-643-2 (1-476) x AF428262 (1-2427)
```

Qy	201	AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr	220
Db	686	GACTGCTGGTCCATTGGTGTTCATCACATACATACATGCTGTGTGGCTATCCCCCTTTCTAT	745
Qy	221	GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer	240
Db	746	GAAAGAACAGAAATCAAGAGCTTTTGAAGATCAAGAGGGGTACTACGAGTTTGAGTCT	805
Qy	241	ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys	260
Db	806	CCGTTCTGGGATGACATTTCTGAGTCAGCCCAAGGATTTATTGGCCATCTGCTGGAGAAG	865
Qy	261	AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn	280
Db	866	GACCCAAACGAAACGGTACACCTGCGGAGAAAGCCCTCAGACACCCCTGGATGTGGAAC	925
Qy	281	ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla	300
Db	926	ACAGCCCTGACCGGGACATCTATCCATCTGTCCAGCCTCCAGATTTCAGAGAACTTTGCC	985
Qy	301	LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu	320
Db	986	AAGAGCAAGTGGAGCAAGCCTTCAACGCAGCTGCGGTGGTCATCATACATGAGGAAACTG	1045
Qy	321	HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr	340
Db	1046	CACATGAACCTGCACAGCCCGCAGTGTCCGCCAAGAGGTGGAGAACAGCCACCTGTGTCC	1105
Qy	341	GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal	360
Db	1106	CCTGCTCTGAAGTCTCCAGACAGACTCCCATTCAGAGCTCCATCAGAGGGCCCCATC	1165
Qy	361	LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg	380
Db	1166	CTGGACCCCAAGCACACACCCCTTCCTGCATGACCCGACTACCTGCTCACACAGCTCCGG	1225
Qy	381	ProThrAlaPro---GlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIle	399
Db	1226	CCTTCAGTCCCAAGTGGTGGCCGCTCACTCAACTGCCCTGGTCAACGGCTCCCTCGCGCATC	1285
Qy	400	SerSerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysCys	419
Db	1286	AGCAGTAGCTAGTGCCTATCAACAGAGGGCCCCCTAGCCAGCGGGCCCTGTGGCTGTGC	1345
Qy	420	SerSerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeu	439
Db	1346	TCCAGCTGTCTAAATATTGGGAACAAGGAAAGTCTTCTACTGCTGTGAGCCTACCCCTC	1405
Qy	440	LeuLysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAla	459
Db	1406	TTCAAGAGGCCAATAAACAACAACTTCAAGTCAGAGGTCAAGTACAGAGGCTGAGAGGCT	1465
Qy	460	SerGlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet	476
Db	1466	GTTGGCAGCAGCCCACTGCGGGGTGGGAGAGTGGGTGTCTCTCGTTATG	1516
RESULT 8			
LOCUS	D86557	Rattus norvegicus mRNA for Protein Kinase, partial cds.	1013 bp mRNA linear ROD 07-FEB-1999
DEFINITION	D86557		
ACCESSION	D86557		
VERSION	D86557.1	GI:2077933	
KEYWORDS	Protein Kinase.		
SOURCE	Rattus norvegicus embryo (E18) brain cDNA to mRNA, clone_lib.S.		
ORGANISM	Nakanishi clone:N5.		
	Rattus norvegicus		
REFERENCE	1 (bases 1 to 1013)		
AUTHORS	Yokokura,H., Terada,O., Naito,Y. and Hidaka,H.		
TITLE	Isolation and comparison of rat cDNAs encoding		

Qy	21	LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal	40
Db	61	AGATCTTCGAGTTCCAAGAGACCCCTCGGAACCGGGCGCTTTCGGAAGGTGGTTAGCT	120
Qy	41	LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle--LysLysSerProAla	59
Db	121	GAAGAGAGGCCAATGGCAAGCTCTTGTGTGGAAGTGTATCCCTAGAAGGGCGCTGAAG	180
Qy	60	PheArgSpSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn	79
Db	181	GCAAGGAAAGCAGCATAGAGAATGAGATGCGCTCTCGAAGAGATTAAAGCATGAAAAAT	240
Qy	80	IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu	99
Db	241	ATTGTGCCCTGGGAAGACATTTATGAAGCCCAATACCTGTACTTGGTCATGCGAGCTG	300
Qy	100	ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp	119
Db	301	GTCTCGGTGGAGAGCTGTGTGACCGGATAGTGGAGAAAGGGGTTTATACAGAGAAAGAT	360
Qy	120	AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle	139
Db	361	GCAGACACTGTATCGCGCAAGCTCTGGAGCGCGGTACTATCTCCACAGAAATGGGCATC	420
Qy	140	ValHisArgSpLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys	159
Db	421	GTCCACAGAGACCTCAAGCCCGAAAAATCTCTGTACTACAGTCAGATGAGGAGTCCAAA	480
Qy	160	IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly--IleMetSerThr	178
Db	481	ATAATGATCAGTGACTTGGATTGTCAAAATGGAGGCAAGGAGATGTGATGCCACT	540
Qy	179	AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys	198
Db	541	GCCTGTGNACTCCAGCCTATGTCGCTCTGAGTCTCGCCAGAAACCTTACACGCANA	600
Qy	199	AlaValAspCysTyrPheIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro	218
Db	601	GCGTGTGACTGCTGTCTCATCGGAGTGATGCTCATCTTGTCTGTGCGGTACCCCTCT	660
Qy	219	PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe	238
Db	661	TTTATGATGAATAGACTCCAAGCTTTTGACAGATCCTCAAGGGCGGAATATGAGTTT	720
Qy	239	GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu	258
Db	721	GACTCTCCCTACTGGGATGACATCTCCGACTTCGAAAAGACTTCATTTCGGAACCTGATG	780
Qy	259	GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrPheLeu	278
Db	781	GAGAGGCCCGCAATAAAGATACACGTGTGACGAGGAGCTCGGCCACCCATGGATCGCT	840
Qy	279	GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn	298
Db	841	GGTGACACAGCCCTCAACAAATAACATCCAGAGTCCGTCCAGCGCCAGCATCCGGAANAAC	900
Qy	299	PheAlaLysSerLysTyrPArgGlnAlaPheAsnAlaAlaValValHisHisMetArg	318
Db	901	TTTGCCACAGACCAATGGAGACAGCATTTAATGCCACGGCGCTCGTGAGACATATGAGA	960
Qy	319	LysLeuHisMet	322
Db	961	AAACTACACCTC	972

[illegible]

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1074)
AUTHORS
Abuho, G., Scoville, J., Turner, C.A., Friedrich, G., Zambrowicz, B.,
Abuin, A. and Sands, A.T.
TITLE
Novel human kinase proteins and polynucleotides encoding the same
JOURNAL
Patent: WO 0142435-A 3 14-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
source
1..1074
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 313 a 256 c 276 g 227 t
ORIGIN

Alignment Scores:
Pred. No.: 3 58e-80 Length: 1074
Score: 1246.00 Matches: 237
Percent Similarity: 85.49% Conservative: 40
Best Local Similarity: 73.15% Mismatches: 45
Query Match: 49.58% Indels: 2
DB: 6 Gaps: 2

```

US-09-960-643-2 (1-476) x AX167587 (1-1074)

QY	20	1	MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg	20
DB	21	1	ATGGCCGGGAGAGCGGCGAGCAGAGCTCCTCTCTGGAAAAAGCAGAGCTGAAGCATCAAG	60
QY	22	21	LysThrPheIlePheMetClnValIleuLysSerClnValAlaPheSerClnValPheLeuVal	40
DB	23	61	ARATCTTCGAAGTCAAAAGACACCTCGGAACCGGGGCCCTTTTCGAAGTGGTTTTAGCT	120
QY	24	41	LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysSerProAla	59
DB	25	121	GAAGAGAAGCAACTGGCAAGCTCTTCTGCTGAAGTGATCCCTAAAGAAGCGCTGAAG	180
QY	26	60	PheArgAspSerSerLeuGluAsnGlnIleAlaValLeuLysLysLysLysHisGluAsn	79
DB	27	181	GGCAAGGAAGCAGCATAGAGAAATGAGATAGCCGTCTCTGAGAAAGATTAAAGCATGA	240
QY	28	80	IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu	99
DB	29	241	ATTGTTGCCCTGGAGACATTTATGAAAGCCCAATACACCTGTACTTGGTCATGCAGCTG	300
QY	30	100	ValSerGlyGlyLeuPheAspArgIleLeuGluArgGlyValTyrThrClnLysAsp	119
DB	31	301	GTCTCCGTGGAGAGCTGTGTGACCGGATATGGAGAAAGGGTTTTATACAGAGAAGAT	360
QY	32	120	AlaSerLeuValIleGlnClnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle	139
DB	33	361	GCCAGCACCTGTATCGCCACAGCTCTTGAGCCGCTGTACTATCTCCACAGAATGGCATC	420
QY	34	140	ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys	159
DB	35	421	GTCCACAGAGACTCAAGCCGCAAAATCTCTGTACTACAGTACAGATGAGAGATGCCAA	480
QY	36	160	IleMetIleThrAspPheGlyLeuSerLysMetClnAlaSnCly---IleMetSerThr	178
DB	37	481	ATAATGATCAGTCACCTTTGGATGTCAAAATGGAGGGCAAGGAGATGTGATGCCACT	540
QY	38	179	AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys	198
DB	39	541	GCCTGTGGAACTCCAGGCTATGTGCTCTCTGAAGTCTCTGCCCGCAAAACCTTACAGCAA	600
QY	40	199	AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro	218
DB	41	601	GCGTTGACTGCGGGTGCACCGAGATGATCCCTACATCTGCTCTGCGGCTACCTCTCT	660
QY	42	219	PheTyrClnLuthrClnSerLysLeuPheGlnLysIleLysGluGlyTyrTyrClnPhe	238
DB	43	661	TTTTATGATGAAATGATCACTCAAGCTCTTTTGAGCAGATCTCTCAAGAGCGGAATATGAGTTT	720

```
Qy 239 GluSerProPheTrpAspPheIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 721 GACTCCCTACTGGATGACATCTCCGACTCTGCAAAAGACTTATTCGGAACTGATG 780

Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpLeuAsp 278
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 781 GAGAAGGACCCGAATAAAGATACAGCTGTGAGCAGGAGCTCGGCACCATGGATCGCT 840

Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 841 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCGGTCAGCCCGCAGATCCGGAANAAC 900

Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 901 TTTGCCAAGACAAATGGAGACAGCATTTAATGCCACGCCGCTCGTGAGACATATGAGA 960

Qy 319 LysLeuHisMet 322
|||||:::|||||:::
Db 961 AACTACACCTC 972

RESULT 11
AX167585
LOCUS AX167585 1158 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 1 from Patent WO0142435.
ACCESSION AX167585
VERSION AX167585.1 GI:14596986
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1158)
AUTHORS Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.,
Abuin,A. and Sands,A.T.
TITLE Novel kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0142435-A 1 14-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1. .1158
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 323 a 283 c 304 g 248 t
ORIGIN

Alignment Scores:
Pred. No.: 3 94e-80 Length: 1158
Score: 1246.00 Matches: 237
Percent Similarity: 85.49% Conservative: 40
Best Local Similarity: 73.15% Mismatches: 45
Query Match: 49.58% Indels: 2
DB: Gaps: 2

US-09-960-643-2 (1-476) x AX167585 (1-1158)

Qy 1 MetGlyArgLysGluGluAspCysSerSerTrpLysGlnThrThrAsnIleArg 20
||| |||::: ||| |||::: ||| |||::: ||| |||::: ||| |||::: ||| |||:::
Db 1 ATGCGCGGAGAACGCGAGACAGCTCTCTCTGGAAGAACGAGCTGAAGACATCAAG 60

Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AAGATCTTCGAGTTCAAAGACCCCTCGGAACCGGGGCCCTTTCCGAAGTGGTTAGCT 120

Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
:::|||||::: ||| |||::: ||| |||::: ||| |||::: ||| |||::: ||| |||:::
Db 121 GAAGAGAAGCAACTGGCAAGCTCTTTGCTGTGAAGTGTATCCCTAAGAAGCGCTGAAG 180

Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
:::|||||::: ||| |||::: ||| |||::: ||| |||::: ||| |||::: ||| |||:::
Db 181 GGCAAGCAAGACAGCATAGAGAATGAGATAGCGCTCTGAGAAAGATTAAAGCATGAAAT 240

Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 ATTGTTCCCTGGAGACATTTATGAAGCCCAATCACCTGTACTTGTGTCATGACGTG 300
```

```
Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 301 GTGTCCCGTGGAGAGCTGTTTGGACCGGATAGTGGAAAGGGGTTTATACAGAAGAGAT 360

Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 361 GCCAGCACTCTGATCCGCCAAGCTCTTGACGCGCTGTACTATCTCCACAGAATGGGCATC 420

Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 421 GTCCACAGAGACCTCAAGCCCGAAATCTCTGTACTACAGTCAAGTACAGAGTCCAAA 480

Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 481 ATAATGATCAGTACTTTGGATTGTCAAAATGGAGGCAAGGAGATGTATGTCCACT 540

Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 541 GCCTGTGAACCTCCAGGCTATGTCGCTCCTGAAGTCTCGCCACAGAACTTACAGCAA 600

Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 601 GCCGTTGACTGCTGGTCCATCGAGTATGTCCTACATCTTGCTCTCGGCTACCCCTCT 660

Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 661 TTTTATGATGAATGACTCCAGCTCTTTGAGCAGATCTCAAGCGCGGATATGATGTTT 720

Qy 239 GluSerProPheTrpAspPheIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 721 GACTCTCCCTACTGGATGACATCTCGGACTCTCGAAAGACTTCAATTCGGAACCTGATG 780

Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 781 GAGAAGGACCCGAATAAAGATACACGCTGTGAGCAGGAGCTCGGCACCATGGATCGCT 840

Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 841 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCCGTCAGCGGCCAGATCCGGAAAAAC 900

Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValAlaValHisMetArg 318
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 901 TTTGCCAAGACAAATGGAGACAGCATTTAATGCCACGCCGCTCGTGAGACATATGAGA 960

Qy 319 LysLeuHisMet 322
|||||:::|||||:::
Db 961 AACTACACCTC 972

RESULT 12
AF286366 1579 bp mRNA linear PRI 17-AUG-2000
LOCUS AF286366
DEFINITION Homo sapiens CamKII-like protein kinase mRNA, complete cds.
ACCESSION AF286366
VERSION AF286366.1 GI:9837340
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1579)
AUTHORS Verploegen,S., Koenderman,L. and Coffey,P.J.
TITLE Identification and characterization of CKIIk: a novel granulocyte
Ca2+/calmodulin-dependent kinase
JOURNAL Blood (2000) In press
REFERENCE 2 (bases 1 to 1579)
AUTHORS Verploegen,S. and Coffey,P.J.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-2000) Dept. Pulmonary Diseases, University
Medical Center Utrecht, Heidelberglaan 100, Utrecht 3584 CX, The
Netherlands
FEATURES
Location/Qualifiers
source 1. .1579
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="eosinophil"
88..1161
/notes="CKL1k"
/codon_start=1
/product="CamK1-like protein kinase"
/protein_id="AAG00534.1"
/db_xref="GI:9837341"
/translation="MARENGSSSWKKOAEIDIKIPFEKTLGTGAPSEVLAERKA
TKLFAVKCIPKALKGKSESIENIAVLRIKHENI VALEDIYESNHLIYLVQLVS
GGELEDRIVEKGYTEKDASTLIRQVLAHYLRHWGIYVHRDLAPENLIYISQDESK
IMISDFGLSKMGKGDVMTAGCTPGYVAPYLAQKPYSKAVDCWSIGVTAYILLCY
PFYDENDSKLEQILKAEYDFSPYDDIIDSADKDFIRNLMEKDPNKRVTCEQAAH
PWAGTALSNKIHESVSAQIRKNFAKSKRQAFNATAVVRHMRKLLHGLSSLDSSNAS
VSSSLSLASOKDAVYVAKPESLS"
BASE COUNT 422 a 414 c 415 g 328 t
ORIGIN
Alignment Scores:
Pred. No.: 5.88e-80 Length: 1579
Score: 1246.00 Matches: 237
Percent Similarity: 85.49% Conservative: 40
Best Local Similarity: 73.15% Mismatches: 45
Query Match: 49.58% Indels: 2
DB: 9 Gaps: 2
US-09-960-643-2 (1-476) x AF286366 (1-1579)
Qy 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysGlnThrThrAsnIleArg 20
Db 88 ATGCCCGGGAGACGCGAGACGAGCTCTCTCTGGAAAAGCAAGCTGAAGACATCAAG 147
Qy 21 LysThrPheIlePheMetGluValValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 148 AAGATCTTCAGATTCAAAAGACACCTCGGAACCGGGCCCTTTCCGAAGTGGTTTAGCT 207
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
Db 208 GAAGAGAAGCAACTGGCAAGCTCTTGTGTGAAGTGTATCCCAAGAAAGCGCTGAAG 267
Qy 60 PheArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 268 GCGAAGAAAGACGACATAGATAGATAGCGCTCTCGAAGAAAGATTGAAGATGAAAT 327
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
Db 328 ATTGTTGCCCTGGAGACATTTATGAAGCCCAATCACCTGTACTTGTCTATGCAGCTG 387
Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
Db 388 GTGTCCCGTGGAGAGCTGTTTGACCGGATAGTGGAGAGGGGTTTWTATACAGAGAAGAT 447
Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValIlyTyrLeuHisGluAsnGlyIle 139
Db 448 GCCAGACTCTGATCCGCCAAGCTCTGGACGCCGTGTACTATCTCCACAGATGGGCATC 507
Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuThrThrProGluGluAsnSerLys 159
Db 508 GTCCACAGACACCTCAAGCCCGAAATCTCTTACTACAGTCAAGATGAGGAGTCCAAA 567
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
Db 568 ATAATGATCAGTGACTTGGATTGTCAAAAATGGAGGGCAAGAGGATGTGATGTCCACT 627
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
Db 628 GCCTGTGGAACTCCAGCGTATGTCGCTCTGAAGTCTCTGCCCCAGAAACCTTACACAAA 687
Qy 199 AlaValAspCysTyrPsrIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 688 GCCCTGTACTCTGCTCATCGGAGTGATTGCCTATACATCTTGTCTCTGCGGCTACCCCTCT 747
```

```
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
Db 748 TTTATGATGAAATGACTCCAGCTCTTTGAGCAGATCTCAAGGCGGAATATGAGTTT 807
Qy 239 GluSerProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 808 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCATTTCGGAACCTGATG 867
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
Db 868 GAGAAGACCCGAAATAAAGATACACGTGTGACGAGGAGCTCGGCACCCATGGATCGCT 927
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
Db 928 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCCGTGAGCGCCAGATCCGGAAGAAC 987
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValAlaValHisMetArg 318
Db 988 TTTCGCCAAGAGCAANTGGAGACAGCATTTAATGCCACGCCGCTGCTGAGACATATGAGA 1047
Qy 319 LysLeuHisMet 322
Db 1048 AAACATACACCTC 1059
RESULT 13
AX167589 LOCUS AX167589 1671 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 5 from Patent WO0142435.
ACCESSION AX167589
VERSION AX167589.1 GI:14596988
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1671)
AUTHORS Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B.,
Abuin,A. and Sands,A.T.
TITLE Novel kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0142435-A 5 14-JUN-2001.
FEATURES
source location/Qualifiers
source 1..1671
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 453 a 437 c 448 g 333 t
ORIGIN
Alignment Scores:
Pred. No.: 6.33e-80 Length: 1671
Score: 1246.00 Matches: 274
Percent Similarity: 64.35% Conservative: 60
Best Local Similarity: 52.79% Mismatches: 97
Query Match: 49.58% Indels: 91
DB: 6 Gaps: 11
US-09-960-643-2 (1-476) x AX167589 (1-1671)
Qy 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 167 ATGCCCGGGAGAACGCGAGACGAGCTCTCTCTGGAAAAGCAAGCTGAAGACATCAAG 226
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 227 AAGATCTTCGAGTTCAAAAGACACCTCGGAACCGGGCCCTTTCCGAAGTGGTTTAGCT 286
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
Db 287 GAAGAGAAGCAACTGGCAAGCTCTTTGCTGTGAGTGTATCCCTAAGAGGCGCTGAAG 346
Qy 60 PheArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 347 GCGAAGAAAGACGACATAGAGAAATGAGATAGCGCTCTCGAAGAAAGATTGAAGCATGAAAT 406
```

```
Qy 80 ILeValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
Db 407 ATTTCTCCCTGGGAAGACATTTATGAAGCCCAATCACCCTGCTACTGTGCTGACGCTG 466
Qy 100 ValSerGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
Db 467 GTGTCGGGTGGAGAGCTGTGTACCGGATAGTGGAGAAGGGTTTATACAGAGAAGAT 526
Qy 120 AlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
Db 527 GCCAGCACTGTATCCGCCAAGTCTTGGAGCGGTGTACTATCTCCACAGAATGGGCATC 586
Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrThrProGluGluAsnSerLys 159
Db 587 GTCCACAGAGACCTCAAGCCGAAAATCTCTGTACAGTCAAGATGAGGAGTCCAAA 646
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
Db 647 ATAATGATCAGTGACCTTGGATTGTCAAAAATGGAGGCAAGGATCTGATGTCCACT 706
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
Db 707 GCCTGTGGAACTCCAGCTATGCTGCTCTGAAGTCTCGCCAGAAACCTTACAGCAA 766
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 767 GCCGTTGACTGCTGGTCCATCGAGTGATGCTTACATCTGCTCGCGCTACCTCTCT 826
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
Db 827 TTTTATGATGAAATGATGCTCAAGCTCTTTGACAGATCTCAAGCGGAAATATGAGTTT 886
Qy 239 GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 887 GACTCTCCCTACTGGGATGACATCTCGACTCTGCAAAAGACTTCATTCGGAACCTGAG 946
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
Db 947 GAGAAGGACCCGAAATAAAGATACAGCTGTGAGCAGGACGCTGGCACCACCTGATCGCT 1006
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
Db 1007 GGTGACACAGCCCTCAACAAAACATCCAGAGTCCGTCAGCGCCCGAGATCCGGAANAAC 1066
Qy 299 PheAlaLysSerIleTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
Db 1067 TTTGCCAAGACAAATGGAGACAAGCATTTAATGCCAGCGCTCGTGAGACATATGAGA 1126
Qy 319 LysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArg----- 336
Db 1127 AAATACACCTCGCGAGCAGCCTGGAGAGTCAAAATGCAAGTGTTCGAGCAGCCTCAGT 1186
Qy 337 ---ProProGluThrGlnAlaSerGluThrSerArgProSerSerProGluIleThrIle 355
Db 1187 TTGCCAGCCCAAAAAGACTGTGCTATGTAGTGAACAAACAGAAATC-CCTCAGCTGACACTG 1245
Qy 356 ThrGluAlaProValLeuAspHisSerValAlaLeuProAla----- 369
Db 1246 AA-GACGAGCTGGGTGGAGAGGAGGAGCGCGGCATCTGCCGAGCAGCCTCCTGTGTGCC 1304
Qy 370 -----LeuThrGlnLeuProCys----- 375
Db 1305 AGCGGCTTTCTATACTTAATCCCATGTCATCGCAGCCCTAGGACTTTTTTTAATCATGTAT 1364
Qy 376 -----GlnHis---GlyArg----- 379
Db 1365 CACTGGGCTGGGTGCAGTGGCTCACGCTGTATATCCCAACACTTCTGGGAGGCTGAGCGAC 1424
Qy 380 -----ArgProThrAlaPro----- 384
Db 1425 GAGGACTGTTGAGTTCAGGAGTTTAAAGACCAAGCTTGACCAACATGTTGAACCCCATC 1484
Qy 385 -----GlyGlyArgSerLeuAsnCysLeuValAsn 394
Db 1485 TCTCTAAATATAAAATTAGCGGTGGTGGCGAGCACCTGTAAATGTCAGCTACTT 1544
Qy 395 GlySerLeuHisIleSerSerSerLeuValProMetHisGlnGlySerLeuAlaAlaGly 414
Db 1545 GGGAGGCTGAGCCA-GGAGAATCGCTTGAGCCC-----GGA 1579
Qy 415 ProCysGlyCysCysSerSerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCys 434
Db 1580 AGCGAGGTTGC-----AATGAGCTGAGATCACACCATGTC 1615
Qy 435 Ser-GluPro-----ThrLeuLeuLysAlaAsnLysLysGln 447
Db 1616 ACTTCAGCTGGGTGACAGATGAGACTCTCTCAAAAAAATAAAAAA 1670
RESULT 14
BC014825 1448 bp mRNA linear ROD 07-AUG-2002
LOCUS BC014825
DEFINITION Mus musculus, Similar to calcium/calmodulin-dependent protein
kinase I, clone MGC:18933 IMAGE:3969343, mRNA, complete cds.
ACCESSION BC014825
VERSION BC014825.1 GI:15928725
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1448)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Liu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 23 Row: j Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
FEATURES
Location/Qualifiers
1..1448
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:18933 IMAGE:3969343"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI_CGAP_Mam1"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
120..1244
/codon_start=1
/product="Similar to calcium/calmodulin-dependent protein
kinase I"
/protein_id="AAH14825.1"
/db_xref="GI:15928726"
CDS
```

/translation="MPGAVEGPRWKAEDIRDIYDFRDVLGTGAFSEVILAEKRTQK
LVAIKCIAKALEGKESMENEIAVLHKIKHPNIVALDDIYESGGHLYLIMOLVSGGE
LFDRIVEKGFYTERDASRLIFQVLDVAKYVLDLGIHVHRLKPNLKYSLDEDSKIMI
SDFGLSKMEDPGSVLSTACGTGCVAPVLAOKPYSKAVDCWISGVIAIYILLCGYPPF
YDENDAKLFEQILKARFEPDSDYDDIISDSAKDFIRHLMKDPKRFCEALQHPWI
AGTALDKNIHQSVSQIKKNFAKSKWQAFNATAVVRHMRKLQLTSGEQGGQTGSH
GELLITPAGGPAAGCCCRDCCVEPSELPPPPSSRAMD"

BASE COUNT 364 a 375 c 408 g 301 t
ORIGIN

Alignment Scores:
Pred. No.: 1,19e-73 Length: 1448
Score: 1157.00 Matches: 237
Percent Similarity: 70.94% Conservative: 51
Best Local Similarity: 58.37% Mismatches: 64
Query Match: 46.04% Indels: 54
DB: 10 Gaps: 8

US-09-960-643-2 (1-476) x BC014825 (1-1448)

Qy 12 TrpLysLysGlnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySer 31
Db 147 TGG---AAGCAGCGGGAAGACATTAGGATATTTATGACTTCAGAGATGTTCTGGCAGC 203
Qy 32 GlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeu 51
Db 204 GTGCTCTTCAGAAAGTATCTTGGCAGGACAGAGGACTCAGAAACTGGTGGCCATC 263
Qy 52 LysCysIle---LysLysSerProAlaPheArgAspSerSerLeuGluAsnGluIleAla 70
Db 264 AAATGCATTGCCAAGAAAGCCCTGGAGGGCCAAAGAGGACGATGCGAGAACGAGATTGCC 323
Qy 71 ValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIleTyrGluSerThr 90
Db 324 GTCTACACAGATCAAGACCCCAACATATTGACCCCTGGATGACATCTATGAGATGGG 383
Qy 91 ThrHisTyrTyrLeuValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIleLeu 110
Db 384 GCCACCTCTACCTCACTGAGTGGTTCAGGTGGAGAGCTGTTGACCGAATTGG 443
Qy 111 GluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnValLeuSerAla 130
Db 444 GAGAAAGGATTCTACAGGAAACGGGATGCGACGGCGCTCATCTCCAGGTGCTGGATGCT 503
Qy 131 ValLysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeu 150
Db 504 GTCAGTACCGACGACGATGGGCTGTCACAGCCGCGATTCGAAGCCAGAGAACCTGCTG 563
Qy 151 TyrLeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMet 170
Db 564 TACTACAGTCTGGATGAAGACTCCAAATCATGATCTTGACITTGCCCTCTCCAGATG 623
Qy 171 GluGlnAsnGly---IleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGlu 189
Db 624 GAGGACCCAGGACGATGGCTCTCCACAGCCTGGGACTCCAGATATGCGCCCTGAG 683
Qy 190 ValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThr 209
Db 684 GTCTGCGCCAGAGCCCTACAGCAGGCTGTGATGCTGTCGTCATAGAGTCATGCTGCC 743
Qy 210 TyrIleLeuLeuCysGlyTyrProPheTyrGluGluThrGluSerLysLeuPheGlu 229
Db 744 TATATAGTCTCTGCTGCTACCCACCCTTTTATGATGATAAATGATGCCAAACTCTTTGAA 803
Qy 230 LysIleLysGluGlyTyrTyrGluPheGluSerProPheTrpAspIleSerGluSer 249
Db 804 CAGATTTTGAAGGACAGATGATGATGACTCTCCCTATTGGGACACATCTCTGACTCT 863
Qy 250 AlaLysAspPheIleCysHisLeuLeuGluLysAspProAsnGluArgTyrThrCysGlu 269
Db 864 GCCAAGATTTTCATACACATTTGATGAGAAAGACCCCGGAGAGAGGTTTACCTGTGAG 923
Qy 270 LysAlaLeuSerHisProTrpIleAspGlyAsnThrAlaLeuHisArgAspIleTyrPro 289

Db 924 CAGGCTTGCAGCACCCATGATGTCAGAGATACACTCTGGATAAGAAATATCCACCAG 983
Qy 290 SerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTyrPArgGlnAlaPheAsn 309
Db 984 TCAGTGACGAGCAGATCAAGAGAACCTTTCGCAAGAGCAGTGAAGCAGCTTTCAT 1043
Qy 310 AlaAlaValValHisHisMetArgLysLeuHisMetAsnLeuHisSerProGly--- 328
Db 1044 GCCACTGTGCTGTGTCGCACATGAGGAAGCTGCGAGCTGGGACCACTCAGGAGGGCAG 1103
Qy 328 ----- 328
Db 1104 GGCAGACTGCGACGCCAGGGGAGCTGCTGACACCAACAGCTGCTGGGCCAGAGTGGC 1163
Qy 329 -----ValArgProGluValGlu-----AsnArgProPro 338
Db 1164 TGCTGCTGCCGAGACTGCTGTGTGGAACCAAGGCTCGGAAGTCCCTGACCAACCCCA 1223
Qy 339 GluThrGlnAlaSerGlu---ThrSerArgProSer----- 349
Db 1224 AGCTCAGGCGCATGATTAACATCAAGGAAGTCAAGGTTTCTTTCATGAGAGGGGTG 1283
Qy 350 -----SerProGluIle-----ThrIleThrGluAla 358
Db 1284 GAGCAGCCTGCTTCTCTCTCCTCCTCAGATCTGGGGGTTCCTGACCCCACTCTTCA 1343
Qy 359 ProValLeuAspHisSer 364
Db 1344 CCCCCTCCCTCTCACTCT 1361
RESULT 15
RATCAMPKAA
LOCUS Rattus norvegicus Cam-like protein kinase mRNA linear ROD 18-MAY-1995
DEFINITION Rattus norvegicus Cam-like protein kinase mRNA, complete cds.
ACCESSION L26288
VERSION L26288.1 GI:439613
KEYWORDS protein kinase
SOURCE Rattus norvegicus (strain Sprague-Dawley) (library: day 21
gestation lung cDNA) fetal lung cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1439)
AUTHORS Cho,F.S., Phillips,K.S., Bogucki,B. and Weaver,T.E.
TITLE Characterization of a rat cDNA clone encoding
calcium/calmodulin-dependent protein kinase I
JOURNAL Biochim. Biophys. Acta 1224 (1), 156-160 (1994)
MEDLINE 95035115
PUBMED 7948038
FEATURES
Location/Qualifiers
1..1439
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="lung"
/dev_stage="fetal"
/tissue_lib="day 21 gestation lung cDNA"
118..1242
/codon_start=1
/product="Cam-like protein kinase"
/protein_id="AAA66944.1"
/db_xref="GI:439614"
/translation="MPGAVEGPRWKAEDIRDIYDFRDVLGTGAFSEVILAEKRTQK
LVAIKCIAKALEGKESMENEIAVLHKIKHPNIVALDDIYESGGHLYLIMOLVSGGE
LFDRIVEKGFYTERDASRLIFQVLDVAKYVLDLGIHVHRLKPNLKYSLDEDSKIMI
SDFGLSKMEDPGSVLSTACGTGCVAPVLAOKPYSKAVDCWISGVIAIYILLCGYPPF
YDENDAKLFEQILKARFEPDSDYDDIISDSAKDFIRHLMKDPKRFCEALQHPWI
AGTALDKNIHQSVSQIKKNFAKSKWQAFNATAVVRHMRKLQLTSGEQGGQTASH
GELLITPAGGPAAGCCCRDCCVEPSELPPPPSSRAMD"
BASE COUNT 357 a 375 c 405 g 302 t
ORIGIN

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 13:47:15 ; Search time 7161 Seconds
(without alignments)
9944.779 Million cell updates/sec

Title: US-09-960-643-1
Perfect score: 2447
Sequence: 1 tggagtggagctcaagcag.....tttttcttaaaaaaaaaa 2447

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	2447	100.0	2447	6	AX399682	Sequence
2	2445.4	99.9	2474	9	BC032787	Homo sapi
3	2415	98.7	2464	9	AF428261	Homo sapi
4	2390	97.7	2612	9	AK095713	Homo sapi
5	1734.8	70.9	1738	9	HS2721161	Homo sapi
6	1328.6	54.3	2416	10	BC021840	Human gen
7	1328.2	54.3	2427	10	AF428262	Mus muscu
8	905.6	37.0	157875	9	HS272116	Human DNA
9	813	33.2	1013	10	D86557	Rattus norv
10	504.6	20.6	1579	9	AF286366	Homo sapi
11	504.6	20.6	1671	6	AX167589	Sequence
12	502.2	20.5	1074	6	AX166517	Sequence
13	502.2	20.5	1074	6	AX167587	Sequence
14	502.2	20.5	1158	6	AX167585	Sequence
15	493	20.1	1439	10	RATCAMPKAA	Sequence
16	489.8	20.0	1480	9	HUMCKI	Homo sapien
17	485	19.8	1402	10	RATPRKI	Rattus norv
18	485	19.8	1448	10	BC014825	Mus muscu
19	428.8	17.5	1618	10	AB004267	Rattus norv
20	428.4	17.5	1032	10	AB023027	Mus muscu
21	428.4	17.5	1332	10	D86556	Rattus norv
22	428.4	17.5	1554	10	AF181984	Mus muscu
23	427	17.4	1032	6	AX166520	Sequence
24	418	17.1	1282	6	AR139102	Sequence
25	291.4	11.9	1546	3	AB021864	Caenorhab
26	290	11.9	3316	3	DMEI7917	Drosophila
27	287.8	11.8	182054	10	AL365314	Mouse DNA
28	287.8	11.8	184754	2	AC022675	Mus muscu
29	250.4	10.2	3141	10	BC017634	Mus muscu
30	247.6	10.1	191584	2	AC120477	Rattus norv
31	245.2	10.0	2897	6	AX099935	Sequence
32	244.4	10.0	3655	10	RATCBVA	Rattus norv
33	243.6	10.0	1503	6	AX239863	Sequence
34	243.6	10.0	2297	6	AX239861	Sequence
35	243.6	10.0	3003	9	BC001921	Homo sapi
36	243.6	10.0	3012	9	BC000497	Homo sapi
37	243.6	10.0	3012	9	BC005828	Homo sapi
38	243.6	10.0	3018	9	BC017363	Homo sapi
39	243.6	10.0	3057	9	BC019256	Homo sapi
40	243.6	10.0	3124	6	AX350342	Sequence
41	229.8	9.4	42404	3	CBRG46J06	Caenorhab
42	214.6	8.8	1701	8	EMECMKA	Emricella
43	197.2	8.1	3471	6	AR027895	Sequence
44	194	7.9	3460	9	HS272212	Homo sapi
45	194	7.9	6612	12	AF416988	Synthetic

ALIGNMENTS

RESULT 1	AX399682	2447 bp	DNA	linear	PAT 06-JUN-2002
LOCUS	Sequence	3 from Patent WO0224947.			
DEFINITION	AX399682				
ACCESSION	AX399682				
VERSION	AX399682.1	GI:21335455			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Delaney,A.D. and Yoganathan,T.				
TITLE	Cancer associated protein kinases and their uses				
JOURNAL	Patent: WO 0224947-A 3 28-MAR-2002;				

KINETEK PHARMACEUTICALS INC (CA); UNIV BRITISH COLUMBIA (CA)	
FEATURES	Location/Qualifiers
source	1..2447
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
CDS	70..1500
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAD33521.1"
	/db_xref="GI:21335456"
	/translation="MGRKEEDDCSSWKQKTNIRKTFIFMEVLGSGAFSEVFLVKQRL
	TKLGLALCIKKSPAFRDSLENEIAVLKKIKHENIVTLEDIYESYTHYLVMOLVSG
	GEFLRILERGVYTKDASLVIOQVLSAVKYLHENGIVHRDLKPENLLYLPMSNSKI
	MITDRGLSKMEONGINSTACGTPGVYAVEVLAQRPYSKAVDCWSIGVTYILLGYP
	FYEETSKLIFEXIKGDIYEFSPFDDIDSEAKDFICHLEKDPNERITCEKALUSHW
	IDENTLHRDIYPSVLIQKNFAKSKRWQAFNAAYVHHMRKLHMLNHSYGVPEVE
	NRPPQASSETSRPSPSEITTEAPVLDHVALPALTOQPOGHRRTTAPGGRSLNCL
	VNGSLHISSSILVPMHQGSLAAGPCSCSLNIGSKGSSYSCEPTLLIKRANKQNF
	SEVMVPVKASGSSHCRAGQTGVCLIM"
BASE COUNT	590 a 707 c 604 g 546 t
ORIGIN	
Query Match 100.0%; Score 2447; DB 6; Length 2447;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 2447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 TGGAGTGGGAGCTCAAGCAGGATCTTCCGAGTCCCTGGCATCCTCAGAACGCTTCAACT 60
Db	1 TGGAGTGGGAGCTCAAGCAGGATCTTCCGAGTCCCTGGCATCCTCAGAACGCTTCAACT 60
Qy	61 CTGAGGCAATGGGTGGAAGGAAGATGACTGCAGTTCCTTGGGAAGAAACAGACCACC 120
Db	61 CTGAGGCAATGGGTGGAAGGAAGATGACTGCAGTTCCTTGGGAAGAAACAGACCACC 120
Qy	121 AACATCCGGAAACCTTCATTTTATGGAAGTGTCTGGGATCAGGAGCTTCTCAGAAGTT 180
Db	121 AACATCCGGAAACCTTCATTTTATGGAAGTGTCTGGGATCAGGAGCTTCTCAGAAGTT 180
Qy	181 TTCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTTTGCTGTGAAGTGCATCAAGAGTCA 240
Db	181 TTCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTTTGCTGTGAAGTGCATCAAGAGTCA 240
Qy	241 CCTGCCCTCCGGACAGCAGCCCTGGAGATGAGATTGCTGTGTGTAAGAAAGATCAAGCAT 300
Db	241 CCTGCCCTCCGGACAGCAGCCCTGGAGATGAGATTGCTGTGTGTAAGAAAGATCAAGCAT 300
Qy	301 GAAACATTTGTACCTTGGAGGACATCTATGAGAGCACCACCACCTACTACTGGTCAATG 360
Db	301 GAAACATTTGTACCTTGGAGGACATCTATGAGAGCACCACCACCTACTACTGGTCAATG 360
Qy	361 CAGCTTCTTGTGGGGAGCTCTTTTGACCGGATCCTGGAGGGGGTGTCTACACAGAG 420
Db	361 CAGCTTCTTGTGGGGAGCTCTTTTGACCGGATCCTGGAGGGGGTGTCTACACAGAG 420
Qy	421 AAGGATCCAGCTGTGTATCCAGAGGCTCTGTCGGCAGTGAATACCTACATGAGAAAT 480
Db	421 AAGGATCCAGCTGTGTATCCAGAGGCTCTGTCGGCAGTGAATACCTACATGAGAAAT 480
Qy	481 GGCACTGTCCACAGAGACTTTAAAGCCCGGAAACCTGCTTTTACCTTACCCTTGAAGAGAC 540
Db	481 GGCACTGTCCACAGAGACTTTAAAGCCCGGAAACCTGCTTTTACCTTACCCTTGAAGAGAC 540
Qy	541 TCTAAGATCATGATCATGCTTGGTCTGTCCAAAGATGGAACAGAAATGGCATCATGTCC 600
Db	541 TCTAAGATCATGATCATGCTTGGTCTGTCCAAAGATGGAACAGAAATGGCATCATGTCC 600
Qy	601 ACTGCCCTGTGGAGCCCGCAGGCTACGTGGCTCCAGAGAGTGTCTGCCCAGAGAACCCCTACAGC 660
Db	601 ACTGCCCTGTGGAGCCCGCAGGCTACGTGGCTCCAGAGAGTGTCTGCCCAGAGAACCCCTACAGC 660
Qy	661 AAGGCTGTGGATTGCTGGTCCATCGGCGTCATCACCTACATATTGCTGTGGATACCCCC 720
Db	661 AAGGCTGTGGATTGCTGGTCCATCGGCGTCATCACCTACATATTGCTGTGGATACCCCC 720

Qy	721 CCCTTCTATGAAGAAACGGAGTCTAAAGCTTTTCGAGAAGATCAAGGAGGCGCTACTATGAG 780
Db	721 CCCTTCTATGAAGAAACGGAGTCTAAAGCTTTTCGAGAAGATCAAGGAGGCGCTACTATGAG 780
Qy	781 TTTGAGTCTCCATTTCTGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTG 840
Db	781 TTTGAGTCTCCATTTCTGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTG 840
Qy	841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTTGAGTCATCCCTGGATT 900
Db	841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTTGAGTCATCCCTGGATT 900
Qy	901 GACGGAACACGGCCCTCCACCGGAGACATCTACCCATCAGTCAGCTCCAGATCCAGAAG 960
Db	901 GACGGAACACGGCCCTCCACCGGAGACATCTACCCATCAGTCAGCTCCAGATCCAGAAG 960
Qy	961 AACTTTTGTAAAGACAAGTGGAGCAAGCCCTTCAACACGACAGCTGTGTGTCACCAATG 1020
Db	961 AACTTTTGTAAAGACAAGTGGAGCAAGCCCTTCAACACGACAGCTGTGTGTCACCAATG 1020
Qy	1021 AGGAAGCTACACATGAACCTTCACAGCCCGGCGTCCGCCAGAGTGGAGAACAGGCCG 1080
Db	1021 AGGAAGCTACACATGAACCTTCACAGCCCGGCGTCCGCCAGAGTGGAGAACAGGCCG 1080
Qy	1081 CCTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTCCCTGTGAGATCACCATCACCGAG 1140
Db	1081 CCTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTCCCTGTGAGATCACCATCACCGAG 1140
Qy	1141 GCACCTGTCTTGGACCAAGTGTAGCACTCCCTGCCCTGACCAATTTACCTGCCAGCAT 1200
Db	1141 GCACCTGTCTTGGACCAAGTGTAGCACTCCCTGCCCTGACCAATTTACCTGCCAGCAT 1200
Qy	1201 GCGCGCCGGCCACTGCCCCCTGGTGGCAGGTCCTCAACTGCCTTGGTCAATGGCTCCCTC 1260
Db	1201 GCGCGCCGGCCACTGCCCCCTGGTGGCAGGTCCTCAACTGCCTTGGTCAATGGCTCCCTC 1260
Qy	1261 CACATCAGCAGAGCCTGGTGCCCATGATCAGGGGTCCCTTGGCCGCGGGCCCTGTGGC 1320
Db	1261 CACATCAGCAGAGCCTGGTGCCCATGATCAGGGGTCCCTTGGCCGCGGGCCCTGTGGC 1320
Qy	1321 TGCTGTCTCAGCTGCTGAACATTTGGGAGCAAGGAAAGTCTCTACTGTCTGTAGCCCC 1380
Db	1321 TGCTGTCTCAGCTGCTGAACATTTGGGAGCAAGGAAAGTCTCTACTGTCTGTAGCCCC 1380
Qy	1381 ACACCTCTCAAAAAGGCGCAACAAAAACAGAACTTCAAGTCGAGAGTCAATGGTACAGTT 1440
Db	1381 ACACCTCTCAAAAAGGCGCAACAAAAACAGAACTTCAAGTCGAGAGTCAATGGTACAGTT 1440
Qy	1441 AAAGCCAGTGGCAGCTCCCACTGCGGCGAGGCGAGCTGGAGTCTGTCTCATTTATGTA 1500
Db	1441 AAAGCCAGTGGCAGCTCCCACTGCGGCGAGGCGAGCTGGAGTCTGTCTCATTTATGTA 1500
Qy	1501 TTCCTGGAGCCTGTGCTTATGTCACTGCAATTTTCAGGAGACATATTCAACTCCTCTGCT 1560
Db	1501 TTCCTGGAGCCTGTGCTTATGTCACTGCAATTTTCAGGAGACATATTCAACTCCTCTGCT 1560
Qy	1561 CTTCCAAACCTGGTGTCTATCCGGCAGAGGAGGAGCAGCAAGTGGAGAGGCGCTT 1620
Db	1561 CTTCCAAACCTGGTGTCTATCCGGCAGAGGAGGAGCAGCAAGTGGAGAGGCGCTT 1620
Qy	1621 AGCAGGACAGTTTCTGGCCAGAGACACAGCTGCTGCCAGCGGGCAGCGCCCTCATAG 1680
Db	1621 AGCAGGACAGTTTCTGGCCAGAGACACAGCTGCTGCCAGCGGGCAGCGCCCTCATAG 1680
Qy	1681 GAGGCCCCAGGAGGAGCCCCAAGGCGTAGAAGCCCTTGTGAAGCTGTGAGCAGGAGAAGC 1740
Db	1681 GAGGCCCCAGGAGGAGCCCCAAGGCGTAGAAGCCCTTGTGAAGCTGTGAGCAGGAGAAGC 1740
Qy	1741 GGTGCCCCACAGCTTCCAGGTCTCCCTGACCTGCTCTCTATGCCCCACACCCCTACGTG 1800
Db	1741 GGTGCCCCACAGCTTCCAGGTCTCCCTGACCTGCTCTCTATGCCCCACACCCCTACGTG 1800

Db 361 CAGCTTGTTTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGGTGTCTACACAGAG 420
QY 421 AAGGATGCCAGTCTGGTGTATCCAGCAGGTCTTTGTGGCAGTGAATACTACATGAGAAT 480
Db 421 AAGGATGCCAGTCTGGTGTATCCAGCAGGTCTTTGTGGCAGTGAATACTACATGAGAAT 480
QY 481 GGCATGCTCCACAGAGACTTAAAGCCGAAACCTGCTTTTACCTTTACCCCTGAAGAGAAC 540
Db 481 GGCATGCTCCACAGAGACTTAAAGCCGAAACCTGCTTTTACCTTTACCCCTGAAGAGAAC 540
QY 541 TCTAAGATCATGATCACTGACTTTGGTCTGTGTCACAGATGGAACAGATGGCATCATGATGCC 600
Db 541 TCTAAGATCATGATCACTGACTTTGGTCTGTGTCACAGATGGAACAGATGGCATCATGATGCC 600
QY 601 ACTGGCTGTGGACCCAGGCTACCTGGCTCCAGAGTGTGGCCAGAAACCTTACAGC 660
Db 601 ACTGGCTGTGGACCCAGGCTACCTGGCTCCAGAGTGTGGCCAGAAACCTTACAGC 660
QY 661 AAGGCTGTGGATGCTGGTCCATCGCGCTCATCACTACATATTCCTGTGGATACCCC 720
Db 661 AAGGCTGTGGATGCTGGTCCATCGCGCTCATCACTACATATTCCTGTGGATACCCC 720
QY 721 CGTCTCTAAGAGAACGGAGTCTAAGCTTTTCGAGAGATCAAGAGGGCTACTATGAG 780
Db 721 CGTCTCTAAGAGAACGGAGTCTAAGCTTTTCGAGAGATCAAGAGGGCTACTATGAG 780
QY 781 TTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTG 840
Db 781 TTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTG 840
QY 841 CTTGAGAAGATCCGAACGAGCGGTACACCTGTGTGAGAGGCTTGTAGTCACTCCCTGGATT 900
Db 841 CTTGAGAAGATCCGAACGAGCGGTACACCTGTGTGAGAGGCTTGTAGTCACTCCCTGGATT 900
QY 901 GACGGAAACACGGCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAG 960
Db 901 GACGGAAACACGGCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAG 960
QY 961 AACTTTGCTAAGAGCAAGTGGAGCAAGCTTTCAACGAGCAGCTGTGGTGCACCATG 1020
Db 961 AACTTTGCTAAGAGCAAGTGGAGCAAGCTTTCAACGAGCAGCTGTGGTGCACCATG 1020
QY 1021 AGGAAGCTACACATGAACTGTACAGCCGGCGGTCCGCCCAGAGGTGGAGAACAGCGG 1080
Db 1021 AGGAAGCTACACATGAACTGTACAGCCGGCGGTCCGCCCAGAGGTGGAGAACAGCGG 1080
QY 1081 CTTGAAACTCAAGCCCTCAGAACTCTAGAACCCAGTCCCTGTAGATCACCATCACCAG 1140
Db 1081 CTTGAAACTCAAGCCCTCAGAACTCTAGAACCCAGTCCCTGTAGATCACCATCACCAG 1140
QY 1141 GCACCTGTCTGTGACACAGTGTAGACATCTCCCTGTGCCCTGACCCAAATACCCCTGCCAGAT 1200
Db 1141 GCACCTGTCTGTGACACAGTGTAGACATCTCCCTGTGCCCTGACCCAAATACCCCTGCCAGAT 1200
QY 1201 GGGCGGGGCCACTGCCCTGGTGGCAGGTCCCTCACTGCTGGTCAATGGCTCCCTC 1260
Db 1201 GGGCGGGGCCACTGCCCTGGTGGCAGGTCCCTCACTGCTGGTCAATGGCTCCCTC 1260
QY 1261 CACATCAGCAGACGCTGTGGCCATGATCAGGGGTCCCTGGCCGCGGGCCCTGTGGC 1320
Db 1261 CACATCAGCAGACGCTGTGGCCATGATCAGGGGTCCCTGGCCGCGGGCCCTGTGGC 1320
QY 1321 TGCTGTCTCCAGTCTCCCACTGGGAGCAAGGAAAGTCTCTCTACTGTCTCTGAGGCC 1380
Db 1321 TGCTGTCTCCAGTCTCCCACTGGGAGCAAGGAAAGTCTCTCTACTGTCTCTGAGGCC 1380
QY 1381 ACACCTCTCAAAAGGCCAACAAAACAGAACTTCAAGTCGGAGGTATGTATACAGTT 1440
Db 1381 ACACCTCTCAAAAGGCCAACAAAACAGAACTTCAAGTCGGAGGTATGTATACAGTT 1440
QY 1441 AAAGCCAGTGGCAGTCCCACTGCCGGCAGGCGAGACTTGGAGTCTGTCTATTATGGA 1500
Db 1441 AAAGCCAGTGGCAGTCCCACTGCCGGCAGGCGAGACTTGGAGTCTGTCTATTATGGA 1500

QY 1501 TTCTTGAGCGCTGCTATGTCTACCTGCAATTTTCAGGAGACATATTCAACTCCTCTGCT 1560
Db 1501 TTCTTGAGCGCTGCTATGTCTACCTGCAATTTTCAGGAGACATATTCAACTCCTCTGCT 1560
QY 1561 CTTCCAAACCTGGTGTCTATCCGCGCAGAGGAGGAGGAGCAAGTGTGAGCAGGCTT 1620
Db 1561 CTTCCAAACCTGGTGTCTATCCGCGCAGAGGAGGAGGAGCAAGTGTGAGCAGGCTT 1620
QY 1621 AGCAGGAGCAGTTTCTGGCCAGAGAACACAGCCTGTGCCAGGGGCGACCCCTCATAG 1680
Db 1621 AGCAGGAGCAGTTTCTGGCCAGAGAACACAGCCTGTGCCAGGGGCGACCCCTCATAG 1680
QY 1681 GAGGCCAGCAGGAGGCCCAAGCGGTAGAACCTTTGTTGAAGCTGTGAGCAGGAGAAC 1740
Db 1681 GAGGCCAGCAGGAGGCCCAAGCGGTAGAACCTTTGTTGAAGCTGTGAGCAGGAGAAC 1740
QY 1741 GGTGCCACACAGCTTCCAGGTCTCCCTGACCTGCTCTATGCCCCACACCTACGTG 1800
Db 1741 GGTGCCACACAGCTTCCAGGTCTCCCTGACCTGCTCTATGCCCCACACCTACGTG 1800
QY 1801 CCGTGGCTCTGTCAGTGTACGTAGATAGCTCTCGCTGGGTCTGTGTTGTCTGTA 1860
Db 1801 CCGTGGCTCTGTCAGTGTACGTAGATAGCTCTCGCTGGGTCTGTGTTGTCTGTA 1860
QY 1861 AAAGCTTTAATGGCTGGCCAGGCTGTACCTTTCTCCAAAGCAAGCCATATGGAGCATC 1920
Db 1861 AAAGCTTTAATGGCTGGCCAGGCTGTACCTTTCTCCAAAGCAAGCCATATGGAGCATC 1920
QY 1921 TACCCAGACTCCCACTCTGCACACACTCACTCCCACTCTCAAGCCTTCCAACTCTTGGC 1980
Db 1921 TACCCAGACTCCCACTCTGCACACACTCACTCCCACTCTCAAGCCTTCCAACTCTTGGC 1980
QY 1981 CAGATTGGGCTCATTAATGTCTGTGGCTGCCCTATCTGCATGAATGACAGCAGCTCCCCA 2040
Db 1981 CAGATTGGGCTCATTAATGTCTGTGGCTGCCCTATCTGCATGAATGACAGCAGCTCCCCA 2040
QY 2041 TGTGGTCTGCTGTGAGCTCTTCAAGTTCTAATCTTAACTCCAGATTAGTCCCAAG 2100
Db 2041 TGTGGTCTGCTGTGAGCTCTTCAAGTTCTAATCTTAACTCCAGATTAGTCCCAAG 2100
QY 2101 TGCTCAGACCCAGCCAGCAGCAGCTTCTGGCCCTTCTCCCTGCTCAATCTAAAGCAGT 2160
Db 2101 TGCTCAGACCCAGCCAGCAGCAGCTTCTGGCCCTTCTCCCTGCTCAATCTAAAGCAGT 2160
QY 2161 GCCACACCTCCAAAGTGAATGAAAGAGTTCATGAGTAAGGGCTGCAAGGAATCTT 2220
Db 2161 GCCACACCTCCAAAGTGAATGAAAGAGTTCATGAGTAAGGGCTGCAAGGAATCTT 2220
QY 2221 ATCTGGCCACATGTCTCTCGTGCACACACCCCAATGGAGTTAACTTGGAGTTGACTAT 2280
Db 2221 ATCTGGCCACATGTCTCTCGTGCACACACCCCAATGGAGTTAACTTGGAGTTGACTAT 2280
QY 2281 TTTAATGTCTGCCAGGAGTTCTAATCTGCTCTGTGTCCCTTTTCTCTCTCTTGAAGTCC 2340
Db 2281 TTTAATGTCTGCCAGGAGTTCTAATCTGCTCTGTGTCCCTTTTCTCTCTCTTGAAGTCC 2340
QY 2341 AGCAGCATTCTGTCTTCCCGAGTTTCTCGCCCTCCACCCCTCCAGCTTTCATGCTC 2400
Db 2341 AGCAGCATTCTGTCTTCCCGAGTTTCTCGCCCTCCACCCCTCCAGCTTTCATGCTC 2400
QY 2401 AGTGTGTCTTAATAAATGGACATATTTTCTCTAAAAA 2447
Db 2401 AGTGTGTCTTAATAAATGGACATATTTTCTCTAAAAA 2447

RESULT 3
AF428261
LOCUS
DEFINITION
ACCESSION
VERSION

AF428261
(CAMK1G) mRNA, complete cds.
AF428261
AF428261.1 GI:16755791

2464 bp
mRNA
linear
PRI 07-NOV-2001
Homo sapiens calcium/calmodulin-dependent protein kinase I gamma

KEYWORDS	Homo sapiens.
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 2464)
AUTHORS	Schutte,B.C., Bjork,B.C., Coppage,K.B., Malik,M.I., Gregory,S.G., Scott,D.J., Brentzell,L.M., Watanabe,Y., Dixon,M.J. and Murray,J.C.
TITLE	A preliminary gene map for the Van der Woude syndrome critical region derived from 900 kb of genomic sequence at 1q32-q41
JOURNAL	Genome Res. 10 (1), 81-94 (2000)
MEDLINE	20113118
PUBMED	10645953
REFERENCE	2 (bases 1 to 2464)
AUTHORS	Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.
TITLE	Characterization of the human ortholog of rat Cam Kinase I gamma (CamK1g) at 1q32-q41
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 2464)
AUTHORS	Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.
TITLE	Direct Submission
JOURNAL	Submitted (05-OCT-2001) Genetics, University of Iowa, 140 EMRB, Iowa City, IA 52242, USA
FEATURES	Location/Qualifiers
source	1..2464
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="1"
	/map="1q32-q41"
gene	1..2464
	/gene="CAMK1g"
CDS	73..1503
	/gene="CAMK1g"
	/codon_start=1
	/product="calcium/calmodulin-dependent protein kinase I gamma"
	/protein_id="AAL28100.1"
	/db_xref="GI:1675792"
	/translation="MGRKEDDSWKKQTNNIKTFIFMEVLGSAFSEVFLVKQRL TKLPAKCIKSPAFROSLSENEIAVLKKIHENIVLEDIYESTHIYLVMLQVSG GEUFLRLRMEGVYTBKDALSVIQVLSAVKYLHENGIVHRDLKPNLDLYLPENSKI MITDEKLMGNGIMSTAGTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLGYPF FYEETSKLFIKEGYEFESFPWDDISEAKDFICHLLEKDPNERTEKALSHPW IDGNTALHRDIYPSVLSQIKFNAFKSKWQAFNAAYVHMRKLHMLNHSIPGRPEVE NRPEPTQASSTRSPSPETITTEAPVLDHSVALPALTPCQHGRRPTAPGRSINCL VNGSLHISSLVPMHQGSLAAGPCGCCSCLNIGNKGSYSSEPTLLKANKQKPK SEVWVPVKASGSSHCRRAGTGVCLIM"
polyA_signal	2415..2420
	/gene="CAMK1g"
BASE COUNT	609 a 708 c 602 g 545 t
ORIGIN	
Query Match	98.7%; Score 2415; DB 9; Length 2464;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 2416; Conservative	0; Mismatches 5; Indels 1; Gaps 1;
Qy	24 TCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACTCTGGAGCAATGGGTCGAAGA 83
Db	27 TCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACTCTGGAGCAATGGGTCGAAGA 86
Qy	84 AGAAGATGACTGAGTTCCTGGGAAGAAACAGACACACATCCGGAAAACCTTCATTTT 143
Db	87 AGAAGATGACTGAGTTCCTGGGAAGAAACAGACACACATCCGGAAAACCTTCATTTT 146
Qy	144 TATGAAGTGTGGATCAGAGCTTCTCAGAAGTTTTCCTGGTGAAGCAAGACTGAC 203
Db	147 TATGAAGTGTGGATCAGAGCTTCTCAGAAGTTTTCCTGGTGAAGCAAGACTGAC 206
Qy	204 TGGGAAGCTCTTTCCTCTGAAGTGCATCAAGAAGTCACTGCTCCCGGACAGCAGCT 263
Db	207 TGGGAAGCTCTTTCCTCTGAAGTGCATCAAGAAGTCACTGCTCCCGGACAGCAGCT 266
Qy	264 GGAGATGAGATTGCTGTGTGAAAAGATCAAGCATGAACATTTGTGACCTGGAGGA 323

Db	267 GGAGATGAGATTGCTGTGTGAAAAGATCAAGCATGAACATTTGTGACCTGGAGGA 326
Qy	324 CATCTATGAGAGACACACACACTACTACCTGGTGCATGCAGCTTGTTCCTGGTGGGAGCT 383
Db	327 CATCTATGAGAGACACACACACTACTACCTGGTGCATGCAGCTTGTTCCTGGTGGGAGCT 386
Qy	384 CTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAGAGAGATGCCAGTCTGGTGATCA 443
Db	387 CTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAGAGAGATGCCAGTCTGGTGATCA 446
Qy	444 GCAGGTCTTGTGGCAGTGAATACCTACATGAGATGGCATGCTCCACAGAGACTTAA 503
Db	447 GCAGGTCTTGTGGCAGTGAATACCTACATGAGATGGCATGCTCCACAGAGACTTAA 506
Qy	504 GCCCGAAAACCTGTTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 563
Db	507 GCCCGAAAACCTGTTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 566
Qy	564 TGGTCTGTCCAGATGAACAGATGGCATGTCTACCTGTCTGTGGAGACCCAGGCTA 623
Db	567 TGGTCTGTCCAGATGAACAGATGGCATGTCTACCTGTCTGTGGAGACCCAGGCTA 626
Qy	624 CGTGGCTCCAGAGTGTGCCCGAGAAACCTTACAGCAAGGCTGTGGATTGTGTGCTCAT 683
Db	627 CGTGGCTCCAGAGTGTGCCCGAGAAACCTTACAGCAAGGCTGTGGATTGTGTGCTCAT 686
Qy	684 CGGCGTCATCACCTTACATATTGTCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTC 743
Db	687 CGGCGTCATCACCTTACATATTGTCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTC 746
Qy	744 TAAGCTTTTCGAGAAGATCAAGAGGCTACTATGATTGAGTCTTCCATCTGGGATGA 803
Db	747 TAAGCTTTTCGAGAAGATCAAGAGGCTACTATGATTGAGTCTTCCATCTGGGATGA 806
Qy	804 CATTTCTGAGTCAGCCAAAGACTTTATTTGCCACTTGTCTGAGAAGATCCGAACGAGCG 863
Db	807 CATTTCTGAGTCAGCCAAAGACTTTATTTGCCACTTGTCTGAGAAGATCCGAACGAGCG 866
Qy	864 GTACACCTGTGAGAAGGCTTGTAGTCACTCCTGGATTGACGGAACACAGCCCTCCACCG 923
Db	867 GTACACCTGTGAGAAGGCTTGTAGTCACTCCTGGATTGACGGAACACAGCCCTCCACCG 926
Qy	924 GGACATCTACCCATCAGTCAGCCTCCAGATCCAGAGAACTTTGTCTAAGAGCAAGTGGAG 983
Db	927 GGACATCTACCCATCAGTCAGCCTCCAGATCCAGAGAACTTTGTCTAAGAGCAAGTGGAG 986
Qy	984 GCAAGCCTTCAACGCGACGCTGTGGTGCACCATGAGGAAGCTACATGAACCTGCA 1043
Db	987 GCAAGCCTTCAACGCGACGCTGTGGTGCACCATGAGGAAGCTACATGAACCTGCA 1046
Qy	1044 CAGCCCGGGCTCCGCCAGAGGTGGAGAACAGCGCCCTGAAACTCAAGCCTCAGAAAC 1103
Db	1047 CAGCCCGGGCTCCGCCAGAGGTGGAGAACAGCGCCCTGAAACTCAAGCCTCAGAAAC 1106
Qy	1104 CTCTAGACCCAGCTCCCTCGATACCATCACCGAGGACCTGTCTCTGGACCAAGTGT 1163
Db	1107 CTCTAGACCCAGCTCCCTCGATACCATCACCGAGGACCTGTCTCTGGACCAAGTGT 1166
Qy	1164 AGCACTCCCTGCCCTGACCCCAATTTACCTGCCAGATGGCGCGGCCCACTGCCCTGG 1223
Db	1167 AGCACTCCCTGCCCTGACCCCAATTTACCTGCCAGATGGCGCGGCCCACTGCCCTGG 1226
Qy	1224 TGGCAGGTCCCTCAACTGCTTGGTCAATGGCTCCCTCCACATCAGCAGACGCTGGTGC 1283
Db	1227 TGGCAGGTCCCTCAACTGCTTGGTCAATGGCTCCCTCCACATCAGCAGACGCTGGTGC 1286
Qy	1284 CATGCATCAGGGGTCCCTTGGCGCGCGGCGCTGTGGCTGTCTGCAGTGCCTGACAT 1343
Db	1287 CATGCATCAGGGGTCCCTTGGCGCGCGGCGCTGTGGCTGTCTGCAGTGCCTGACAT 1346
Qy	1344 TGGGAGCAAGGAAGTCTCTACTGCTGTGAGCCCACTCTCTCAAAAAGGCCAACAA 1403

```
Db 1347 TGGGAACAAAGGAAGTCCCTCTACTGCTCTGAGCCCAACACCTCTCTCAAAAAGGCCAACAA 1406
Qy 1404 AAAACAGAACTTCAAGTCGAGGTCTATGGTACCAGTAAAGCCAGTGGCAGTCCCACTG 1463
Db 1407 AAACAGAACTTCAAGTCGAGGTCTATGGTACCAGTAAAGCCAGTGGCAGTCCCACTG 1466
Qy 1464 CCGGGCAGGCGAGACTGGAGTGTCTCTCAATATGTATGTCTCTGGAGCCCTGTGCCTATGTC 1523
Db 1467 CCGGGCAGGCGAGACTGGAGTGTCTCTCAATATGTATGTCTCTGGAGCCCTGTGCCTATGTC 1526
Qy 1524 ACTGCAATTTTCAGGAGACATATCAACTCTCTGCTCTTCCAAACCTGGTGTCTATCG 1583
Db 1527 ACTGCAATTTTCAGGAGACATATCAACTCTCTGCTCTTCCAAACCTGGTGTCTATCG 1586
Qy 1584 GCAGAGGAGGAGGCGAGCAAGTGGAGCAGGCTTAGCAGAGCAGTTCCTGGCCAGA 1643
Db 1587 GCAGAGGAGGAGGCGAGCAAGTGGAGCAGGCTTAGCAGAGCAGTTCCTGGCCAGA 1646
Qy 1644 AGCACCAAGCTGTGCGAGCGGGCAGCCCTCATAGGAGGCCAGGAGGGAGGCCCAAG 1703
Db 1647 AGCACCAAGCTGTGCGAGCGGGCAGCCCTCATAGGAGGCCAGGAGGGAGGCCCAAG 1706
Qy 1704 GCGTAGAAGCTTGTGAAGCTGTGAGCAGAGAGAGGGTGGCCACACAGTCTCCAGGTCT 1763
Db 1707 GCGTAGAAGCTTGTGAAGCTGTGAGCAGAGAGAGGGTGGCCACACAGTCTCCAGGTCT 1766
Qy 1764 CCCTGAGCTGCTCTATGCCCCACACCCCTACGTGCCGTGCTGTGCGAGTGTACGT 1823
Db 1767 CCCTGAGCTGCTCTATGCCCCACA - CCTAGCTGCCGTGCTGTGCGAGTGTACGT 1825
Qy 1824 AGATAGTCTGCGCTGGGTCTGTGCTGTTTGTGCGTGAAGCTTAATGGCTGGCCAGCG 1883
Db 1826 AGATAGTCTGCGCTGGGTCTGTGCTGTTTGTGCGTGAAGCTTAATGGCTGGCCAGCG 1885
Qy 1884 TGTGTCACTTCTCCAAAGCAAGCATATGAGCAGTCTACCCAGACTCCCACTCTGCACA 1943
Db 1886 TGTGTCACTTCTCCAAAGCAAGCATATGAGCAGTCTACCCAGACTCCCACTCTGCACA 1945
Qy 1944 CACTCACTCCCACTCTCAAGCCTTCCAACTCTTGGCCAGATTGGCTCATTAATGTCT 2003
Db 1946 CACTCACTCCCACTCTCAAGCCTTCCAACTCTTGGCCAGATTGGCTCATTAATGTCT 2005
Qy 2004 TGCCTGCCCTCTGCATGAATGACAGGCACCTCCCATGGTGTCTGCTGTGAGTCTCT 2063
Db 2006 TGCCTGCCCTCTGCATGAATGACAGGCACCTCCCATGGTGTCTGCTGTGAGTCTCT 2065
Qy 2064 CAAGTCTTAATCCCTTAAGTCCAGGATTAGCTCCCAAGTGGCTGAGACCCAGCCAGCA 2123
Db 2066 CAAGTCTTAATCCCTTAAGTCCAGGATTAGCTCCCAAGTGGCTGAGACCCAGCCAGCA 2125
Qy 2124 CTTCGTGCCCTTCTCCCTGCTCAATCTAAAGCAGTGGCCACACCTCCCAAGTGAATA 2183
Db 2126 CTTCGTGCCCTTCTCCCTGCTCAATCTAAAGCAGTGGCCACACCTCCCAAGTGAATA 2185
Qy 2184 GAAAGAGTTCATGAGTAAAGGCTGCAAGCAATTTATCTGCGCCACATGCTCCCGG 2243
Db 2186 GAAAGAGTTCATGAGTAAAGGCTGCAAGCAATTTATCTTCTGCGCCACATGCTCCCGG 2245
Qy 2244 CACACACCAATGAGTAACTTTGGAAGTTGACTATTTTAATGTCTGCCAGGAGTTCTA 2303
Db 2246 CACACACCAATGAGTAACTTTGGAAGTTGACTATTTTAATGTCTGCCAGGAGTTCTA 2305
Qy 2304 ATCCTGCTCTGTTCCTTTCTCTCTGAAAGTCCAGACACATTTCTGTCTCTCC 2363
Db 2306 ATCCTGCTCTGTTCCTTTCTCTCTGAAAGTCCAGACACATTTCTGTCTCTCC 2365
Qy 2364 CAGTTTCTGCTCCCTCCACCCCTCCAGCTTTCATGCTCAGTGTGTGCTTAATAAATGGA 2423
Db 2366 CAGTTTCTGCTCCCTCCACCCCTCCAGCTTTCATGCTCAGTGTGTGCTTAATAAATGGA 2425
Qy 2424 CATATTTTCTCTAAAAA 2447
Db 2426 CATATTTTCTCTAAAAA 2449
```

```
RESULT 4
LOCUS AK095713
DEFINITION Homo sapiens cDNA FLJ38394 fis, clone FEBRA2007534, moderately similar to CamK1-like protein kinase.
ACCESSION AK095713
VERSION Oligo capping: fis (full insert sequence).
KEYWORDS Homo sapiens fetal brain cDNA to mRNA, clone_lib:FEBRA2
SOURCE clone:FEBRA2007534.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M.,
            Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
            NEDO human cDNA sequencing project
            Unpublished
            2 (bases 1 to 2612)
REFERENCE Isogai,T. and Yamamoto,J.
            Direct Submission
            Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 252-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
            NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES
            Location/Qualifiers
                1..2612
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="FEBRA2007534"
                /tissue_type="brain"
                /clone_lib="FEBRA2"
                /dev_stage="fetal"
                /note="cloning vector: pME18SFL3"
BASE COUNT 619 a 750 c 645 g 598 t
ORIGIN
Query Match 97.7%; Score 2390; DB 9; Length 2612;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2393; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 39 GCATPCTCTCAGAGCTTCAACTCTCGAGGCAATGGTTCGAAGGAAGAGATGACTGCAG 98
Db 215 GGCATCTCTCAGAGCTTCAACTCTCGAGGCAATGGTTCGAAGGAAGAGATGACTGCAG 274
Qy 99 TTCTCTGGAAGAACAGACACCAACATCCGGAACCTTCATTTTATGGAAGTCTGTGG 158
Db 275 TTCCTTGAAGAACAGACACCAACATCCGGAACCTTCATTTTATGGAAGTCTGTGG 334
Qy 159 ATCAGAGCTTTCCTCAGAAGTTTTCCTGGTGAAGCAAAAGACTGACTGGGAAGCTTTTC 218
Db 335 ATCAGAGCTTTCCTCAGAAGTTTTCCTGGTGAAGCAAAAGACTGACTGGGAAGCTTTTC 394
Qy 219 TCTGAAGTGCATCAGAGTACCTGCTCCGGACAGCAGCCCTGGAGATGAGATGTC 278
Db 395 TCTGAAGTGCATCAGAGTACCTGCTCCGGACAGCAGCCCTGGAGATGAGATGTC 454
Qy 279 TGTGTTGAAAAGATCAAGCATGAAACATTTGTACCCTGGAGGACATCTATGAGAGCAC 338
```

|||||
Db 455 TGTGTTGAAAAAGATCAAGCATGAAAAATTTGTGACCCCTGGAGGACATCTATGAGAGCAC 514
QY 339 CACCCACTACTACCTGGTCATGACAGCTGTGTTCTGGTGGGAGCTCTTTGACCCGATCCT 398
Db 515 CACCCACTACTACCTGGTCATGACAGCTGTGTTCTGGTGGGAGCTCTTTGACCCGATCCT 574
QY 399 GGAGCGGGTGCTACACAGAGAAGGATGCCAGTGTGCTGTGATCCACAGCAGGCTCTTGTGGGC 458
Db 575 GGAGCGGGTGCTACACAGAGAAGGATGCCAGTGTGCTGTGATCCACAGCAGGCTCTTGTGGGC 634
QY 459 AGTGAATACCTACATGAGATGGCATCGTCCACAGAGACTTAAAGCCCGAAAACCTTGCT 518
Db 635 AGTGAATACCTACATGAGATGGCATCGTCCACAGAGACTTAAAGCCCGAAAACCCGCT 694
QY 519 TTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTTTGCTGTGCTCAAGAT 578
Db 695 TTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTTTGCTGTGCTCAAGAT 754
QY 579 GGAACAGATGSCATCATGTCCACTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAAT 638
Db 755 GGAACAGATGSCATCATGTCCACTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAAT 814
QY 639 GCTGGCCCAAGAAACCTACAGCAAGGCTGTGGATTGCTGTGCTCCATCGGCGTCATCACCTA 698
Db 815 GCTGGCCCAAGAAACCTACAGCAAGGCTGTGGATTGCTGTGCTCCATCGGCGTCATCACCTA 874
QY 699 CATATTGCTGTGGATACCCGCCCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAA 758
Db 875 CATATTGCTGTGGATACCCGCCCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAA 934
QY 759 GATCAAGGAGGCTACTATGACTTTGAGTCTCCATTCTGGGATGACATTTCTGAGTCAGC 818
Db 935 GATCAAGGAGGCTACTATGACTTTGAGTCTCCATTCTGGGATGACATTTCTGAGTCAGC 994
QY 819 CAAAGCACTTTATTGGCCACTTCTTTGAGAAGGATCCGACAGCGGCTACACCTGTGAGAA 878
Db 995 CAAAGCACTTTATTGGCCACTTCTTTGAGAAGGATCCGACAGCGGCTACACCTGTGAGAA 1054
QY 879 GGCCTTGACTCATCCCTGGATTGAGGAAACAGCGCCCTCCACCGGACATCTACCCATC 938
Db 1055 GGCCTTGACTCATCCCTGGATTGAGGAAACAGCGCCCTCCACCGGACATCTACCCATC 1114
QY 939 AGTCAGCCTCCAGATCCAGAGAAGCTTTGCTAAGAGCAAGTGGAGGCAAGCCTTCAACGC 998
Db 1115 AGTCAGCCTCCAGATCCAGAGAAGCTTTGCTAAGAGCAAGTGGAGGCAAGCCTTCAACGC 1174
QY 999 AGCAGCTGGTGCCACCATGAGGAGCTACACATGAACTGACACGCCCGGGGCTCCG 1058
Db 1175 AGCAGCTGGTGCCACCATGAGGAGCTACACATGAACTGACACGCCCGGGGCTCCG 1234
QY 1059 CCCAGAGGTGGAGAACAGAGCCGCTGAAACTCAAGCCTCAGAAACCTCTAGACCCAGCTC 1118
Db 1235 CCCAGAGGTGGAGAACAGAGCCGCTGAACTCAAGCCTCAGAAACCTCTAGACCCAGCTC 1294
QY 1119 CCCTGAGATACCATCAGCGAGGACCTGTCTCTGGACACAGTGTAGCACTTCCCTGGCCCT 1178
Db 1295 CCCTGAGATCAGCATCAGCGAGGACCTGTCTCTGGACACAGTGTAGCACTTCCCTGGCCCT 1354
QY 1179 GACCCAAATTACCTGCCAGCATGGCGGCCCTGAACTCAAGCCTCAGAAACCTCTAGACCCAGCTC 1238
Db 1355 GACCCAAATTACCTGCCAGCATGGCGGCCCTGAACTCAAGCCTCAGAAACCTCTAGACCCAGCTC 1414
QY 1239 CTGCCCTGGTCAATGGCTCCCTCCACATCAGCAGCAGCCTGGTGGCCCATGCAATCAGGGGTC 1298
Db 1415 CTGCCCTGGTCAATGGCTCCCTCCACATCAGCAGCAGCCTGGTGGCCCATGCAATCAGGGGTC 1474
QY 1299 CCTGGCCCGGGCCCTGTGGCTGTGCTCCAGCTGCCCTGAACATTTGGAGCAAGGAA 1358
Db 1475 CCTGGCCCGGGCCCTGTGGCTGTGCTCCAGCTGCCCTGAACATTTGGAGCAAGGAA 1534
QY 1359 GTCTCTCTACTGCTGTGAGCCACACTCCTCAAAAAGGCCAACAAAAACAGAACTTCAA 1418
|||||

RESULT 5
HS272L161

Db 1535 GTCTCTCTACTGCTGTGAGCCACACTCCTCAAAAAGGCCAACAAAAACAGAACTTCAA 1594
QY 1419 GTCGAGGTCATGTTACCAAGTTAAAGCCOAGTGGCAGTCCCACCTGCCGGCAGGCGAC 1478
Db 1595 GTCGAGGTCATGTTACCAAGTTAAAGCCOAGTGGCAGTCCCACCTGCCGGCAGGCGAC 1654
QY 1479 TGGAGTCTGTCTCATTTATGTGATTCCTGGAGCCTGTGCCCTATGTCTACCTGCAATTTTCAGG 1538
Db 1655 TGGAGTCTGTCTCATTTATGTGATTCCTGGAGCCTGTGCCCTATGTCTACCTGCAATTTTCAGG 1714
QY 1539 AGACATATTCAACTCTCTGCTCTTCCAAACCTGGTGTCTATCCGGCAGAGGGAAGG 1598
Db 1715 AGACATATTCAACTCTCTGCTCTTCCAAACCTGGTGTCTATCCGGCAGAGGGAAGG 1774
QY 1599 CAGAGCAAGTGGAGCAGGCTTAGCAGGAGCAGTTTCTGGCCAGAGCACCAGCCTGTGTG 1658
Db 1775 CAGAGCAAGTGGAGCAGGCTTAGCAGGAGCAGTTTCTGGCCAGAGCACCAGCCGCTG 1834
QY 1659 CAGCCGGGAGGCCCTCATAGGAGGCCAGGAGGAGGCCCAAGCGTAGAAGCTTGT 1718
Db 1835 CCAGCGGGCAGGCCCTCATAGGAGGCCAGGAGGAGGCCCAAGCGTAGAAGCTTGT 1894
QY 1719 TGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGCCTGC 1778
Db 1895 TGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGCCTGC 1954
QY 1779 TCTATGCCCCACACCTACTGTCGCGTGTGTGTCAGTGTACGTAGATGATCTCTCGCT 1838
Db 1955 TCTATGCCCCACACCTACTGTCGCGTGTGTGTCAGTGTACGTAGATGATCTCTCGCT 2014
QY 1839 GGTCTGTGCTGTCTGCTGTAAGAGCTTAATGGGCTGCCAGGCTGTGTACCTTCTCC 1898
Db 2015 GGTCTGTGCTGTCTGCTGTAAGAGCTTAATGGGCTGCCAGGCTGTGTACCTTCTCC 2074
QY 1899 AAGCAAGCCATATGAGCATCTACCCAGACTCCCACTCTGCACACACTCACTCCCACT 1958
Db 2075 AAGCAAGCCATATGAGCATCTACCCAGACTCCCACTCTGCACACACTCACTCCCACT 2134
QY 1959 CTCAAGCTCCCACTCTTGGCAGATTTGGGCTCATTAATGCTGTGGCTGCCATCTGC 2018
Db 2135 CTCAAGCTCCCACTCTTGGCAGATTTGGGCTCATTAATGCTGTGGCTGCCATCTGC 2194
QY 2019 ATGAATGACAGCAGCTCCCACTGTGTGTCTGCTGTGAGCTCTTCAAGTCTTAATCTCT 2078
Db 2195 ATGAATGACAGCAGCTCCCACTGTGTGTCTGCTGTGAGCTCTTCAAGTCTTAATCTCT 2254
QY 2079 AACTCCAGATTAAGCTCCCACTGTGTGTCTGCTGTGAGCTCTTCAAGTCTTAATCTCT 2138
Db 2255 AACTCCAGATTAAGCTCCCACTGTGTGTCTGCTGTGAGCTCTTCAAGTCTTAATCTCT 2314
QY 2139 CTTGCTCTCAATCTAAAGCAGTGGCAGCCTCCCAAGTGGAAATAGAAAGAGTTCATGA 2198
Db 2315 CTTGCTCTCAATCTAAAGCAGTGGCAGCCTCCCAAGTGGAAATAGAAAGAGTTCATGA 2374
QY 2199 GTAAGGCTGCAAGGAATTTATCTCTGCCACATGCTCTCCGTGGACACACCCCAATGA 2258
Db 2375 GTAAGGCTGCAAGGAATTTATCTCTGCCACATGCTCTCCGTGGACACACCCCAATGA 2434
QY 2259 GTTAACCTTGGAAATGAGCTATTTTAAATGCTGCCAGGAGTTCTAATCCTGCTGTGTC 2318
Db 2435 GTTAACCTTGGAAATGAGCTATTTTAAATGCTGCCAGGAGTTCTAATCCTGCTGTGTC 2494
QY 2319 CCTTTTCTCTCTGAAAGTCCAGCACCAATTTGTCTCTTCCCCAGTTTCTCTGCCCT 2378
Db 2495 CCTTTTCTCTCTGAAAGTCCAGCACCAATTTGTCTCTTCCCCAGTTTCTCTGCCCT 2554
QY 2379 CCACCCCTCCAGCTTCTGCTGAGTGTGTGCTTAATAAATGGACATATTTTCTCT 2436
Db 2555 CCACCCCTCCAGCTTCTGCTGAGTGTGTGCTTAATAAATGGACATATTTTCTCT 2612


```
QY 1073 ACAGGCGCGCTGAACCTCAAGCCTCAGAACCTCTAGAACCCAGCTCCCTCGATGATCACCA 1132
Db 1021 ACAGGCGCGCTGAACCTCAAGCCTCAGAACCTCTAGAACCCAGCTCCCTCGATGATCACCA 1080
QY 1133 TCACCGAGGACCTGCTCCCTGGACACAGTGTAGCACTCCCTGCCCTGACCCCAATTAACCT 1192
Db 1081 TCACCGAGGACCTGCTCCCTGGACACAGTGTAGCACTCCCTGCCCTGACCCCAATTAACCT 1140
QY 1193 GCAGAGATGCCCGCGCCCACTGCGCCCTGGTGGCAGTCCCTCAACTGCCCTGGTCAATG 1252
Db 1141 GCCAGATGCCCGCGCCCACTGCGCCCTGGTGGCAGTCCCTCAACTGCCCTGGTCAATG 1200
QY 1253 GCTCCCTCCACATCAGCAGCAGCTGTGCCATCATCAGGGTCCCTGGCGCGCGGCG 1312
Db 1201 GCTCCCTCCACATCAGCAGCAGCTGTGCCATCATCAGGGTCCCTGGCGCGCGGCG 1260
QY 1313 CTTGTGGCTGCTCTCCAGCTGCTGAACATTTGGGAGCAAGAAAGTCTCTCTACTGCT 1372
Db 1261 CTTGTGGCTGCTCTCCAGCTGCTGAACATTTGGGAGCAAGAAAGTCTCTCTACTGCT 1320
QY 1373 CTGAGCCCACTCTCAAAAAGGCCAACAAAACAGAACTTCAAGTCGGAGGTCATGG 1432
Db 1321 CTGAGCCCACTCTCAAAAAGGCCAACAAAACAGAACTTCAAGTCGGAGGTCATGG 1380
QY 1433 TACAGTTAAAGCCAGTGGCAGCTCCCTGCTGCTGCGGCGGCGAGACTGGAGTCTGCTCA 1492
Db 1381 TACAGTTAAAGCCAGTGGCAGCTCCCTGCTGCTGCGGCGGCGAGACTGGAGTCTGCTCA 1440
QY 1493 TTATGTGATTCCTGGAGCCTGTGCTATGCTCATCTGCAATTTTCAGGAGACATATCAACT 1552
Db 1441 TTATGTGATTCCTGGAGCCTGTGCTATGCTCATCTGCAATTTTCAGGAGACATATCAACT 1500
QY 1553 CTTCTGCTCTTCAAACTGGTGTCTATCCGCGAGGAGGAGAAAGCAGAGCAATGGAG 1612
Db 1501 CTTCTGCTCTTCAAACTGGTGTCTATCCGCGAGGAGGAGAAAGCAGAGCAATGGAG 1560
QY 1613 CAGGGCTTAGCAGGACAGCTTTCTGGCCAGAAAGCAGCCTGCTGCGAGGGGCGAGCC 1672
Db 1561 CAGGGCTTAGCAGGACAGCTTTCTGGCCAGAAAGCAGCCTGCTGCGAGGGGCGAGCC 1620
QY 1673 CCTCATAGAGGCGCCAGGAGGAGGCCCAAGCGGTAGAGCCTTGTGAAGCTGTGAGCA 1732
Db 1621 CCTCATAGAGGCGCCAGGAGGAGGCCCAAGCGGTAGAGCCTTGTGAAGCTGTGAGCA 1680
QY 1733 GGAGAACGGTGCACACAGCTTCAGAGTCTCCCTGACCTGCTGCTATGCCCCAC 1790
Db 1681 GGAGAACGGTGCACACAGCTTCAGAGTCTCCCTGACCTGCTGCTATGCCCCAC 1738

RESULT 6
BC021840 2416 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, clone MGC:30513 IMAGE:4502479, mRNA, complete cds.
DEFINITION BC021840
ACCESSION BC021840.1 GI:18256866
VERSION BC021840.1
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.
Direct Submission
Submitted (18-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 41 Row: 1 Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: Genomescan gene
prediction.

FEATURES

source	Location/Qualifiers
1..2416	
/organism="Mus musculus"	
/db_xref="taxon:10090"	
/clone="MGC:30513 IMAGE:4502479"	
/tissue_type="Eye, retina, mouse strain C57Bl/6"	
/clone_lib="NIH_MGC_94"	
/lab_host="DH10B"	
/note="Vector: pCMV-SPORT6"	
87..1520	
/codon_start=1	
/product="Unknown (protein for MGC:30513)"	
/protein_id="AAH21840.1"	
/db_xref="GI:18256867"	
/translation="MGKKEEDCSNWKQKTTNIRKTFIPMEVLGSCAPSEVFLVKQV TGKLPALCKIKSPAPRDSLENEIAVLARKIHENITVLEDIESTHTYLYMQLVSG GELFDRILRGVYTKDASLVQQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKI MITDFGLSKMEQNGVSTAGCTPGYVAPVLAQKPKYKADKWCSTGYITILLGYP FYETESKLFKEIKGYEFGSPFDDISAKDFICHLEKDPNERYTCERALKRHPW IDGNTALHRDLYPSVLSIOKFNKAFKRWQAFNAAVVHMKLHMNLHSPSVROEVE NRPPVSPAPEVSRPDSITFAIPILDPSTPLPALTRLPCSHSRPSAPSGRSLNC LVNGLSRLSSSLVPMOOGPLATGPGCCSCSLNIGNKSSYCSBETLPRKANKQNF KSEVWPVRAAGSTHCRGGGTGCLVM"	
BASE COUNT 572 a 694 c 602 g 548 t	

ORIGIN	Query Match	Score	DB 10;	Length	2416;
	Best Local Similarity	76.1%;	Pred. No. 0;		
	Matches 1853;	Conservative	0;	Mismatches	494;
				Indels	88;
				Gaps	14;
QY 31	GAGTCCCTGGCATCCTCAGAGCTTCAACTCTGGAGCAATGGTGCAGAAAGAGAGAT	90			
Db 48	GAGCCCTGGCTTCTCAGGAGCTTCAACTCTGGAGCATGGGCGTAAGGAGGAGGAG	107			
QY 91	GACTGCAGTTCCTGGAAGAACAGACACCACCAACATCGGAAACCTTCATTTATGGAA	150			
Db 108	GACTGCAGTTCCTGGAAGAACAGACACCACCAACATCGGAAACCTTCATTTATGGAA	167			
QY 151	GTGCTGGGATCAGAGCTTCTCAGAAAGTTCCTGGTGAAGCAAAAGACTGACTGGGAAG	210			
Db 168	GTGCTGGGATCAGAGCTTCTCAGAGTGTCTCCTGGTGAAGCAAAAGACTGACTGGGAAG	227			
QY 211	CTCTTTGCTCTGAAGTGCATCAGAGTCACTACCTGCTCCGCGACAGCAGCCTGGAGAAT	270			
Db 228	CTCTTTGCTCTGAAGTGCATCAGAGTCACTACCTGCTCCGCGACAGCAGCCTGGAGAAT	287			
QY 271	GAGATTCCTGTTGAAAGATCAAGCATGGAACATTTGACCCCTGGAGGACATCTAT	330			
Db 288	GAGATTCCTGTTGAAAGATCAAGCATGGAACATTTGACCCCTGGAGGACATCTAT	347			
QY 331	GAGAGCACCCCACTACTACCTGGTGCATCAGCTGTTCTTGTGGGAGCTCTTTGAC	390			
Db 348	GAGAGCACCCCACTACTACCTGGTGCATCAGCTGTTCTTGTGGGAGCTCTTTGAC	407			
QY 391	CGATCTCGAGCGGGGTGTCTACACAGAGAAGATGCCAGTCTGGTGATCCAGAGGTC	450			
Db 408	CGATCTCGAGCGGGGTGTCTACACAGAGAAGATGCCAGTCTGGTGATCCAGAGGTC	467			

Qy	451	TTGTGCGCAGTGAATACCTACATGAGAATGSGCATCGTCCACAGAGACTTTAAAGCCCGAA	510
Db	451		
Db	468	TTGTGCGGTGAATACCTTCATGAGAATGSCATCGTCCACAGAGATCTAAAGCCTGAA	527
Qy	511	AACCTGCTTTACCCTTACCCCTTGAAGAGAACCTTAGATCATGATCACTAGCTTTGGTCTG	570
Db	511		
Db	528	AACCTGCTGTACCTTCAACCCCTGAGAGAACTCCAAGATCATGATCACTAGCTTTGGTCTA	587
Qy	571	TCCAAGATGGAACAGAATGGCATCATGTCCACTGCGCTGGGAGCCCAAGGCTACGTGGCT	630
Db	571		
Db	588	TCCAAGATGGAGAGAATGGAGTCATGTCCACAGCTTGTGGAGCCCAAGGCTACGTGGCT	647
Qy	631	CCAGAAGTGTGGCCAGAAAACCTTACAGCAAGGCTGTGGATTTGCTTGGTCCATCGGGCTC	690
Db	631		
Db	648	CCAGAAGTGTGGCCAGAAAGCCCTACAGTAAAGGCTGTGGACTGCTGGTCCATTTGGTGT	707
Qy	691	ATCACCTTACATATTTGCTGTGGATACACCCCGTTCTTATGAAGAAAGGAGCTCTAAGCTT	750
Db	691		
Db	708	ATCACATACATCTGCTGTGTGGCTATCCCTCTTCTATGAAGAAACAGAATCAAGCTT	767
Qy	751	TTCGAGAAAGTCAAGGAGGGCTACTATGAGTTTGAGTCTCCATTTCTGGGATGACATTCT	810
Db	751		
Db	768	TTTGAGAAGATCAAGAGGGTTACTACGAGTTTGAGTCTCCGTTCTGGGATGACATTCT	827
Qy	811	GAGTCAGCCAAAGACTTTATTTGCCACTTTGCTTGAGAAGGATCCGAACAGAGCGGTACACC	870
Db	811		
Db	828	GAGTCAGCCAAAGATTTATTTGCCATCTGCTGGAGAAGGACCCAAACGAACGGGTACACC	887
Qy	871	TGTCAGAAAGCCTTGAGTCATCCCTGGATTGACGGAACACGGGCCCTCCACCGGGACATC	930
Db	871		
Db	888	TGGAGAAAGCCCTCAGACACCCCTGGATTGATGAAGAACACAGCCCTTGCACCGGGAATC	947
Qy	931	TACCCATCAGTCAGCCTCCAGATCCAGAGAAGAACTTTGCTTAAGACCAAGTGGAGGCAAGCC	990
Db	931		
Db	948	TATCCATCTGTACGCTCCAGATTCAAGAAGAACTTTTGCCAAAGCAAGTGGAGGCAAGCC	1007
Qy	991	TTCAAGCGCACGCTGTGTGCACACATGAGGAAGCTTACACATGAACCTGCACAGCGCCG	1050
Db	991		
Db	1008	TTCAACGACGCTCGGTGTGTGCATCACAATGAGGAACCTGCACATGAACCTGCACAGCGCC	1067
Qy	1051	GGGTCCGCCACAGAGTGGAGAACAGCGCCCTGAAACTCAAGCCTCAGAAACCTCTAGA	1110
Db	1051		
Db	1068	AGTGTCCGCCAAGAGTGGAGAACAGGCCACCTGTGTCCCTGTCTGAAGTCTCCAGA	1127
Qy	1111	CCNAGCTCCCTTGAGATCAGCATCACCAGGAGCACCTGTCTTGAGCCACAGCTGTAGCACTC	1170
Db	1111		
Db	1128	CCAGACTCCCATACAGACTCCATCACAAGAGGCCCCCATCTCTGGACCCCAAGCACACCCCTT	1187
Qy	1171	CCTGCCCTGACCCAATTACCCCTGCACAGATGGCCCGCGGCCCATCTGCCCC---TGGTGGC	1227
Db	1171		
Db	1188	CTGCACTGACCCGACTACCCCTGCTCACACAGCTCCCGGCCCTTCAGCTCCCAAGTGGTGGC	1247
Qy	1228	AGTGCCTCAACTGCCTGTGTCAATGGCTCCCTCCACATCAGCAGGAGCGCTGGTGCCTCATG	1287
Db	1228		
Db	1248	CGCTCACTCAACTGCTGTGTCAACGGCTCCCTTGCACATCAGCAGTAGCTAGTGCCTCATG	1307
Qy	1288	CATCAGGGTTCCTGTGCGCGCGGGCCCTGTGGCTGCTGCTCCAGCTGCTCAACATTTGGG	1347
Db	1288		
Db	1308	CAACAGGGCCCTTAGCCAGGGGCCCTGTGGCTGCTGCTCCAGCTGTCTAAATATTGGG	1367
Qy	1348	AGCAAGAGAAAGTTCCTCTACTGCTGTGAGCCCAACATCTCTCAAAAAGGCCCAACAAAA	1407
Db	1348		
Db	1368	AACAAGGGAAGCTTCTCTACTGCTGAGCTACCCCTCTTCAGAAAGGCCCAATAAAAA	1427
Qy	1408	CAGAACCTTCAAGTCGGAGTCAATGTFACCAGTTAAAGCCAGTGGCAGCTCCCACTGCCGG	1467
Db	1408		
Db	1428	CAGAACTTCAAGTCAGAGGTATGTTACAGTGAAGGCTGTGGCAGCACCCACTGCCGG	1487
Qy	1468	GCAGGCGAGACTGGAGTCTGTCTCAATTATGTGATTTCTTGGAGCCTTGTGCCTATGTCACTG	1527
Db	1468		
Db	1488	GGTGGGCGAGACTGGGTGTGTCTGCTTATGTGATCATCCAGAGGCCCATGTGGTTTTTCAGGA	1547
Db	1488		

QY	1528	CAATTTTCAGGAGACATATTTCAACTCTCTGTCTCTTCCAAACCTCGTGTCTATCTGGCAG	1587
DB	1548	GACATTCCTG--GCTCTTCTCTCCCTCTCTGAGCTGGCATCTGCCCTCAGGAAGGGGGC	1605
QY	1588	AGGAGGAAGGCAGAGCAAGTGCAGCAGGCGCTTTAGCAGGAGCAGTTTCTGCGCCAGCA	1647
DB	1606	AGTGGAAATAGAGATAGCCAAAGCAGCTGGTCTGAGGCAG-----GCCCACCAGCT	1644
QY	1648	CCAGCCTGCTGCCAGCGGGCAGCCCTCTATAGGAGGCCCCAGGAGGAGCCCCAAGCGGT	1707
DB	1645	TAGGCGCTGTTGCTGCCCTGGCAGACACTCACAGAGAACTCCAGAGGAGCGACGCCAATCT	1704
QY	1708	AGAAGCCTTGTTGAAGCTGTGACGAGGAGAGCGGT-----GCCCACCAGCT	1754
DB	1705	TGGACGTTCTCTGAGGCCACAAGCAGGATGAGTGGCGCTGCCCTCCAGCATCCAGCC	1764
QY	1755	TCCAGGCTCCCTGACCTGCGTCTCTATGCCCCACACCCCTACGTCGCGTGGCTGTGTGC	1814
DB	1765	TCCAGGCTGCC--GACCTGGCCATTGCCCTCCCCACATCATACGTGTCAGTGGCTCCATGC	1823
QY	1815	AGTGTAGCTAGATAGCTTCGCGCTGGTCTGTGCTGTTTGCCTGCTGAAAGCTTAATGGC	1874
DB	1824	AGTGTCCATAGATAGTCTCATCTGGGTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC	1883
QY	1875	TGGCAGGCGTGTGTCACTTCTCCAAAGCAAAAGCCATATGGAGCATCTACCCAGACTCCCA	1934
DB	1884	TGGCCAGGCTGTGTCACTTTCTTCAAGCAAAAGCCATATGGAGTCTTAACCTCAGACTCCCTG	1943
QY	1935	CTCTGACACACTCACTCCACACTCTCAAGCCTCCAAACCTCTTGGCCAGATATGGGCTCAT	1994
DB	1944	CTCTGACAGCTCGGCTCTGCTCATCT-----GGCCTCAGGCCAGACTGGACCCCT	1994
QY	1995	TAATGTGCTTGGCTGCCCATCTGCATGAATGACAGCAGCAGCTCCCCATGGTGTGCTGCGTG	2054
DB	1995	TATCATAG--TAGCTGCGCTATTGCGCATGAATGCCAGGCTGCACC-----AGGGGCGCTG	2045
QY	2055	TGAGCTCTTCAAGTTCTAATCCTTAACCTCAGGATTAGCTCCCAAGTGGCTGGCTGAGACCCA	2114
DB	2046	GGAGCATCTACAAGTCCAGTCCGTTCTCCAGGATTAGCTTCCCCTGCAGCGAGACGCCA	2105
QY	2115	GCCAGCACACTTCTGGCGCTTCTCCCTGCGCTCAATCTAAAGCAGTG--CCACACCCCTCCA	2173
DB	2106	GGCCACTC-----CAAAAGCTCTCTTGGCTCCATCTCAGCAGTGCCGCCCATCTCCA	2159
QY	2174	AAGTGGAAATAGAAGAAGTTCATGATAGGGCTGCAAGGAATCTTATCTGGGCCACAT	2233
DB	2160	TGGTGGAGTACAGAAAGCTGTGACTATGGAGCTGTGAGCTTCTCCCTGCTGGGCCACAC	2219
QY	2234	GTCTCTCGTGACACACCCCAATGGAGTTAACTTGGAACTTGACATATTTAATGTCTGCC	2293
DB	2220	ACCTCATATATCCCTAATTAAAG-----ATTGGGGACTATTTCATGTGTCATCC	2268
QY	2294	AGGAGTTCTAATCTGCTCTGTGTCCCTTTCTCTCCTTGAAGTCCAGCAGCACCATCTCT	2353
DB	2269	AGG-----TCTCTGCTCTGCTCCCTCCCTTCCCTCCCTGCTGGGCCACAC	2322
QY	2354	TGCTCTTCCCGAGTTTCTCTCGCCCTCCACCCCTCCAGCTTCATGCTCAGTGTGTGCTCT	2412
DB	2323	-----GTCTCTCTGTGTCTTGCTCATACCTCTCAAGCTCTGCTGTGTGTGTGCTCA	2377
QY	2413	AATAAAATGGACATATTTTTTCTCTAAAAAAA	2447
DB	2378	AATAAAATGGACATATTTTTTCTCTAAAAAAA	2412

RESULT 7
AF428262
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

AF428262 2427 bp mRNA linear ROD 07-NOV-2001
Mus musculus calcium/calmodulin-dependent protein kinase I gamma
(Camk1g) mRNA, complete cds.
AF428262
AF428262.1 GI:16755793

Db 1487 GGTGGCAGAGCTGGGTGTCTCGTTATGTGATCCAGGAGCCCATGTGGTTTTCAGGA 1546

QY 1528 CAATTTTCAGGAGACATATTAACACTCTCTGCTCTTCCAAACCTGGTGTCTATCCGGGAC 1587

Db 1547 GACATTCTCTGCTCTCTGCTCTCTGAGCTGGCATCTGCCCTGAGGAAGGGGGG 1604

QY 1588 AGGAGGAGGAGCAGACAGTGGAGCGGCTTAGCAGGAGCAGTTCTTCGCGCCAGAAGCA 1647

Db 1605 AGTGGAGATAGAGATAGCAAGCAGTGTGTCTGAGGCAG----- 1643

QY 1648 CCAGCCTGTCTGCCAGGGGAGCCCTCATAGGAGGCCAGGAGGAGCCCAAGGCGT 1707

Db 1644 TAGGGCTGTGTCTGGCGCCACACTCACAGGAAGTCCAGAGGAGCCCAACTCT 1703

QY 1708 AGAAGCCTGTTCAGCTGTGAGCAGGAGGAGGCT-----GCCACCAAGCT 1754

Db 1704 TGGACGTCTCTGAGGCCCAAGCAGAGTGTGGGCTGCCCTCCAGCATCCAGCC 1763

QY 1755 TCCAGGTCTCCCTGAGCTGCTCTATGCCCCACACCTACGTGCCCTGTGTCTGTGC 1814

Db 1764 TCCAGGTCTCCC-GACCTGCCATTCCTTCCCCACATCATAGTGCAGTGCCTCCATGC 1822

QY 1815 AGTGTAGTAGATAGTCTCGCTGGCTGTGCTGTGCTGTGCTGTAAGCTTAATGGC 1874

Db 1823 AGTGTCCATAGATAGTCTCATCTGGGCTCGTGTGTTGTGTGAAAGCTTTATGGC 1882

QY 1875 TGGCCAGGCTGTCTACCTTCTCCAGCAAGCCATATGGAGCATCTACCCAGACTCCCA 1934

Db 1883 TGGCCAGGCTGTCTACTTCTTCAGCAAGCCATATGAGTCTTAACCTCAGACTCCCTG 1942

QY 1935 CTCTGCACACACTCTCCACCTCTCAAGCCTCCAACTCTTGGCCAGATTGGGCTCAT 1994

Db 1943 CTCTGCAGACTCTCTCTGCTCTATCT-----GGCCTCAGCGCCAGACTGGACCCAT 1993

QY 1995 TAATGCTGTGCTGCTCCATCTCCATGATGACAGCAGCTCCCCATGCTGTGCTGCTG 2054

Db 1994 TATCATAG-TACCTGTGCTATTCGAGTGTGCGAGTGTGCGAGCTGCACC-----AGGGGCGCTG 2044

QY 2055 TGAGCTCTCAAGTTCTTAATCTTAACTCCAGGATTAGCTCCCAAGTGGCTGAGACCCA 2114

Db 2045 GGAGCACTCACAACTCCAGTCCGTTCTCCAGGATTAGCTTCCCACTGCACGAGACCA 2104

QY 2115 GCGAGCACACTCTGCGCCCTTCTCCCTGCTCAATCAATCAAGAGTGC-CCACACCCCTCA 2173

Db 2105 GGCACCTC-----CAAAGCTCTCTTGTGCTCCATCTCAGCAGTGTGCCCCCATCTCCA 2158

QY 2174 AAGTGGATAGAAAGAGTTCATGAGTAAAGGCTGCAAGGAATCTTATCTTGGCCACAT 2233

Db 2159 TGGTGGAGTATAGAAAGCTGTGACTATGAGCTGTGAGCTTCTCCCTGCTGGCCACAC 2218

QY 2234 GTCTCTGCTGCACACACCCCAATGGAGTTAACTTTGGAAGTTGACTATTATATGTGCTG 2293

Db 2219 ACCTCATATATCCCTAATTAAG-----ATTGCGGGACTATTTCATGTGATCC 2267

QY 2294 AGGAGTCTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2353

Db 2268 AGG-----TCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2321

QY 2354 TGTCTCTTCCCAAGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2412

Db 2322 -----GTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2376

QY 2413 AATAAATGACATATTTTCT 2447

Db 2377 AATAAATGACATATTTTCT 2411

RESULT 8

LOCUS HS272L16 157875 bp DNA linear PRI 23-NOV-1999

DEFINITION Human DNA sequence from clone 272L16 on chromosome 1q32.1-32.3. Contains the 3' end of the LAMB3 gene for Laminin, Beta 3 (Nilein, Kalinin, BM600) and a novel Rat Ca2+/Calmodulin dependent Protein

Kinase LIKE gene. Contains ESTs, STSS, GSSs, genomic marker D1S491 and a ca repeat polymorphism, complete sequence.

AL023754 GI:4007152

VERSION 1

KEYWORDS HFG; BM600; ca repeat polymorphism; Ca2+/Calmodulin dependent Protein Kinase; D1S491; Kalinin; LAMB3; Laminin Beta 3; Nilein.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 157875)

AUTHORS Grafham, D.

TITLE Direct Submission

JOURNAL Submitted (27-NOV-1998) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

COMMENTS requests: clonerequest@sanger.ac.uk

On Dec 12, 1998 this sequence version replaced gi:3873472. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence is the entire insert of clone 272L16. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

272L16 is from the library RPCII constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:pcypac2> This sequence was generated from a human chromosome 1 bacterial clone contig constructed in collaboration by the Sanger Centre chromosome 1 mapping group and Brian Schutte, Bryan Bjork, Kevin Coppage and Jeffrey Murray. Department of Pediatrics, University of Iowa, USA. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>.

FEATURES

Location/Qualifiers

1..157875

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/map="q32.1-32.3"

/clone="RP1-272L16"

/lib="RP1-1"

966..1078

/note="MLT11 repeat: matches 281..410 of consensus"

1082..1506

/note="MSTD repeat: matches 1..426 of consensus"

1585..1875

/note="MLT1A1 repeat: matches 1..319 of consensus"

1898..2024

/note="MLT11 repeat: matches 58..203 of consensus"

2347..2925

/note="MER34 repeat: matches 6..543 of consensus"

3090..3417

/note="MLT1A2 repeat: matches 23..374 of consensus"

3677..4074

/note="MER54B repeat: matches 486..902 of consensus"

4089..5338

/note="MLT1A2-internal repeat: matches 358..1643 of consensus"

5339..5532

/note="AluJo repeat: matches 101..302 of consensus"

5536..5966

/note="MLT1A1-internal repeat: matches 5..450 of consensus"

5972..6019

/note="MLT1A1 repeat: matches 318..365 of consensus"

6020..6459

/note="LTR7 repeat: matches 1..450 of consensus"

Db	147711	CAGCTTCCAGGTCCTCCCTGACCTGCCTGCTCTATGCCCCACACCCCTAGCTGCCGTGGCTC	147770
QY	1810	TGTGCAGTGTACGTAGATAGCTCTCGCCTGGGCTCTGTGCTGTTTCTGCTGTAAGACTTAA	1869
Db	147771	TT	
Db	147771	TGTGCAGTGTACGTAGATAGCTCTCGCCTGGGCTCTGTGCTGTTTCTGCTGTAAGACTTAA	147830
QY	1870	TGGGCTGGCCAGGCTGTGCACCTTCTCCAAGCAAGCCATATGGAGCATCTACCCAGAC	1929
Db	147831	TGGGCTGGCCAGGCTGTGCACCTTCTCCAAGCAAGCCATATGGAGCATCTACCCAGAC	147890
QY	1930	TCCCACTGTGCACACACTCACTCCCACTCTCAAGCTCCCAACTCTTGGCCAGATTGGG	1989
Db	147891	TCCCACTGTGCACACACTCACTCCCACTCTCAAGCTCCCAACTCTTGGCCAGATTGGG	147950
QY	1990	CTCATTAAATGCTGTGCTGCCCATCTGCATGAATGACAGGAGCTGCCCATGGTGGTCT	2049
Db	147951	CTCATTAAATGCTGTGCTGCCCATCTGCATGAATGACAGGAGCTGCCCATGGTGGTCT	148010
QY	2050	GCCTGTGAGCTCTTCAAGTTCTAATCCTTAACTTCCAGGATTAGCTCCCAAGTGCCTGAG	2109
Db	148011	GCCTGTGAGCTCTTCAAGTTCTAATCCTTAACTTCCAGGATTAGCTCCCAAGTGCCTGAG	148070
QY	2110	ACCCAGCCAGCACACTTCTGGCCCTTCTCCCTGCCTCAATCTAAAGAGAGTGGCACACCC	2169
Db	148071	ACCCAGCCAGCACACTTCTGGCCCTTCTCCCTGCCTCAATCTAAAGAGAGTGGCACACCC	148130
QY	2170	TCCAAAGTGAATAGAAAGAGTTGATGAGTAAGGGCTCAAGGAAATCTTATCTCTGGCC	2229
Db	148131	TCCAAAGTGAATAGAAAGAGTTGATGAGTAAGGGCTCAAGGAAATCTTATCTCTGGCC	148190
QY	2230	ACATGTCTCTCCGTGCACACACCAATGGAGTTAACTTGGAAAGTTGACTATTTTAAATGTC	2289
Db	148191	ACATGTCTCTCCGTGCACACACCAATGGAGTTAACTTGGAAAGTTGACTATTTTAAATGTC	148250
QY	2290	TGCCAGGAGTTCTAATCCTGCTGCTCTGTTCCCTTTTCTCTCTCTCTCTCTCTCTCTCTCT	2349
Db	148251	TGCCAGGAGTTCTAATCCTGCTGCTCTGTTCCCTTTTCTCTCTCTCTCTCTCTCTCTCTCT	148310
QY	2350	TTCTTGTCTCTTCCCAAGTTTCTCTGCGCCCTCCACCCCTCCAGCTTTCATGCTCAGTGTGTG	2409
Db	148311	TTCTTGTCTCTTCCCAAGTTTCTCTGCGCCCTCCACCCCTCCAGCTTTCATGCTCAGTGTGTG	148370
QY	2410	CTTAATAAATGGACATATTTTCTCTTA	2437
Db	148371	CTTAATAAATGGACATATTTTCTCTTA	148398
RESULT 9			
LOCUS	D86557	1013 bp	mRNA linear
DEFINITION	Rattus norvegicus mRNA for Protein Kinase, partial cds.		
ACCESSION	D86557		
VERSION	D86557.1	GI:2077933	
KEYWORDS	protein kinase.		
SOURCE	Rattus norvegicus embryo (E18) brain cdna to mRNA, clone_lib:S. Nakanishi clone:N5.		
ORGANISM	Rattus norvegicus		
REFERENCE	1 (bases 1 to 1013)		
AUTHORS	Yokokura,H., Terada,O., Naito,Y. and Hidaka,H.		
TITLE	Isolation and comparison of rat cDNAs encoding		
JOURNAL	Ca2+/calmodulin-dependent protein kinase I isoforms		
MEDLINE	Biochim. Biophys. Acta 1338 (1), 8-12 (1997)		
REFERENCE	97228532		
AUTHORS	2 (bases 1 to 1013)		
TITLE	Yokokura,H.		
JOURNAL	Submitted (15-JUL-1996) Hisayuki Yokokura, Nagoya University School		
TITLE	of Medicine, Department of Pharmacology; Tsurumai 65, Showa-ku,		
JOURNAL	Nagoya, Aichi 466, Japan (Tel:052-744-2075, Fax:052-744-2083)		
FEATURES	Location/Qualifiers		

source	1. .1013		
	/organism="Rattus norvegicus"		
	/db_xref="taxon:10116"		
	/clone="N5"		
	/tissue_type="brain"		
	/clone_lib="S. Nakanishi"		
	/dev_stage="embryo (E18)"		
5'UTR	<1. .86		
CDS	87. .>1013		
	/note="similar to calcium/calmodulin-dependent protein		
	kinase I mRNA (L24907, L26288)"		
	/codon_start=1		
	/product="Protein Kinase"		
	/protein_id="BAA19880.1"		
	/db_xref="GI:2077934"		
	/translation="MGRKEEDCSWKQTTNIRKTFIRMEVLGSGAFSEFLVKQRV		
	TGLFALKCIKSPAFRDSLENIASVAVKYLKRIHENIVLTLEDIYESTHYVLYMLVSG		
	GELFDRILERTKENDSLVIOQVLSAVKYLHENGIVHRDLKAPENLLLTPEENSKI		
	MITDFGLSKMEQNGVSTACGTGYVAPEVLAQKPYSKAVDCWSIGVITLILCGTTP		
	FYEETSKLFKEIKEGYIFEFSPWDDISAKDFICHLEKDPNERYTCEKALRHPW		
	IDGNTALHRDIYPSVSLQIQKNFAKSKWRQAFN"		
BASE COUNT	276 a 245 c 262 g 230 t		
ORIGIN			
	Query Match 33.2%; Score 813; DB 10; Length 1013;		
	Best Local Similarity 90.2%; Pred. No. 1.1e-213;		
	Matches 870; Conservative 0; Mismatches 95; Indels 0; Gaps 0;		
QY	31	GAGTCCCTGGCATCCTCAGAGCTTCAACCTCGGAGCAATGGTCCGAAAGGAGAGAT	90
Db	48	GAGCCCTGGCTTCCCAGGCACTTCAACCTCGGAGGACATGGGGCGAAAGGAGGAG	107
QY	91	GACTGCACTTCTCGAAGAAACAGACCCCAACATCCGGAAACCTTCATTTTATGAA	150
Db	108	GACTGTAGCTCTGGAAGAACACACCCCAACATCAGGAAACCTTCATCTTCATGAA	167
QY	151	GTGCTGGGATCAGGAGCTTCTCAGAAGTTTTCCTGGTGAAGCAACACTGACTGGGA	210
Db	168	GTGCTGGGATCAGGAGGCTTCTCAGAGGTGTCTCGTGAAGCAAGAGTGAAGGAAA	227
QY	211	CTCTTTGCTCTGAAGTGCATCAAGAAAGTGCACCTCCCTCCGGGACACAGCTGGAGA	270
Db	228	CTCTTTGCTCTGAAGTGCATCAAGAAAGTGCACCTCCCTCCGGGACAGTAGAGA	287
QY	271	GAGATTCGTGTGTTGAAAGAGTCAAGCAATGAAAAACATTTGACCCCTGGAGGACAT	330
Db	288	GAGATTCGTGTGTTGAAAGAGTCAAGCAATGAAAAACATTTGACCCCTGGAGGAT	347
QY	331	GAGAGCACCCCACTACTACCTGTCATGCAGCTTCTTTCTGTTGGTGGAGCTCTTTGAC	390
Db	348	GAGAGCACCCCACTACTACCTGTCATGCAGCTTCTTTCTGTTGGTGGAGCTTTTTCAC	407
QY	391	CGGATCCTCGAGCGGGTGTCTACACAGAGAAGGATGCCAGTCTGTGTGATCCAGAGTC	450
Db	408	CGAATCCTAGAGCGTGGTGTCTACACAGAAAAGGATGCCAGTCTGTGTATCCAGAGTC	467
QY	451	TTGTCGGGAGTGAATACCTACATGAGAAATGGCATCTGCCACAGAGACTTTAAAGCCGAA	510
Db	468	TTGTCGGGAGTGAATACCTACATGAGAAATGGCATCTGCCACAGAGACTTTAAAGCCGAA	527
QY	511	AACCTGCTTACCTTACCCCTGAAGAGAATCTTAAGATCATGATCATGACTTTGGTCTG	570
Db	528	AACCTGCTTACCTTACCCCTGAAGAGAATCTTAAGATCATGATCATGACTTTGGTCTG	587
QY	571	TCCAAGATGGAACAGAAATGGCATGATGTCCTGCTGCTGGGAGCCCGAGGCTAGTGGCT	630
Db	588	TCCAAGATGGAACAGAAATGGCATGATGTCCTGCTGCTGGGAGCCCGAGGCTAGTGGCT	647
QY	631	CCAGAGTGTGTCGCCCAAGAAACCTTACAGAGAGCTGTGGATTCTGTCATCCATCGGCTC	690
Db	648	CCAGAGTGTGTCGCCCAAGAAACCTTACAGAGAGCTGTGGATTCTGTCATCCATCGGCTC	707
QY	691	ATCACCTACATATTGCTCTCTGTTGGATACCCCGCTTCTATGAAGAAACGAGGACTTAAGCTT	750

KEYWORDS	human.
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1671)
AUTHORS	Donoho,G., Scoville,J., Turner,C.A., Friedrich,G., Zambrowicz,B., Abuln,A. and Sands,A.T.
TITLE	Novel human kinase proteins and polynucleotides encoding the same
JOURNAL	Patent: WO 0142435-A 5 14-JUN-2001;
FEATURES	Lexicon Genetics Incorporated (US)
source	Location/Qualifiers
	1..1671
BASE COUNT	453 a 437 c 448 g 333 t
ORIGIN	
	Query Match 20.6%; Score 504.6; DB 6; Length 1671;
	Best Local Similarity 71.3%; Pred. No. 3.3e-128;
	Matches 695; Conservative 0; Mismatches 274; Indels 6; Gaps 2;
Qy	66 GGCAATGGGTGGAAAGGAGAGATGACTGCAGTTCTTGGAAAGAACAGACCACCAACAT 125
Db	
Qy	126 CCGGAAACCTTCATTTTATGGAAGTCTCGGATCAGGAGCTTCTCAGAAGTTTTCCT 185
Db	
Qy	185 CCGGAAACCTTCATTTTATGGAAGTCTCGGATCAGGAGCTTCTCAGAAGTTTTCCT 185
Db	
Qy	223 CAAGAAGATCTTCGAGTTCAAGAGACCTCGGAACCGGGGCCCTTTCCGAAGTGGTTT 282
Db	
Qy	186 GGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATC --- AAGAAGTCAC 242
Db	
Qy	283 AGCTGAAGAGAAGGCAACTGGCAAGCTCTTTGCTGTGAAGTATCCTTAAGAAGCGCT 342
Db	
Qy	243 TGCTTCCGGGACAGCAGCTGGAGATGAGATTGCTGTGTGTAAGAAATCAAGCATGA 302
Db	
Qy	302 AAACATTTGACCCCTGGAGGACATCTATGAGAGCACCACCACTACTACTGCTGCTATGCA 362
Db	
Qy	403 AAATATTTGGCCCTGGAAGACATTTATGAAGCCCAAAATCACCTGACTTGGTCATGCA 462
Db	
Qy	363 GCTTGTCTTGGTGGGAGCTCTTGGACCGATCCTGGAGGGGGTGCTACACAGAA 422
Db	
Qy	463 GCTGGTCTCGGTGGAGAGCTGTTTGACCGATAGTGGGAAGGGGTTTATACAGAA 522
Db	
Qy	423 GGATGCCAGTCTGGTGATCCACAGGTCTCTCGGCAGTGAATAACCTACATCAGAAATGG 482
Db	
Qy	523 GGATGCCAGCACTCTGATCCGCCAGTCTTGAGCGCGTGTACTATCTCCACAGATGG 582
Db	
Qy	483 CATCTGCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAACTC 542
Db	
Qy	583 CATCTGCCACAGAGACTCAAGCCGAAATCTCTTGTACTACAGTCAAGATCAGAGATC 642
Db	
Qy	543 TAAGATCATGATCACTGACTTTGGTCTCCAAAGATGGAACAGAAATGG --- CATCATGTC 599
Db	
Qy	643 CAAATATATGATCAGTACTTTGGATGTCAAAATGGAGGCAAGAGAGATGTGATGTC 702
Db	
Qy	600 CACTGCTCTGGGACCCAGGCTAGTGGCTCCAGAAAGTGTCTGCCACAGAAACCTACAG 659
Db	
Qy	703 CACTGCTCTGGAACTCCAGGCTATGTCGCTCCTGAAGTCTCGCCACAGAAACCTTACAG 762
Db	
Qy	660 CAAGGCTGTGGATTGCTGGTCCATCGGGCTCATCCTACATATGCTCTGGGATACCC 719
Db	
Qy	763 CAAAGCCGTTGATCTGCTGGTCCATCGGAGTGAITTCCTACATCTTGCTCTGGGGCTACCC 822
Db	
Qy	720 CCGTCTTATGAAGAAACGGAGTCTFAAGCTTTTTCGAGAAGATCAAGGAGGCTTACTATGA 779
Db	
Qy	823 TCCTTTTATGATGAATAGTCCAGCTCTTTGACAGATCCTCAAGCGGGAATATGA 882
Db	
Qy	780 GTTTGAGTCTCCATCTCTGGGATGACATTTCTGATCGACCAAGACTTTATTTGCCACTT 839
Db	
Qy	883 GTTTGAGTCTCCATCTCTGGGATGACATCTCGACTCTGCAAAAGACTTCAATTCGGAACCT 942
Db	

Qy	840	GCTTGAGAAGGATTCGAACGAGCGGTGCATCCTGTGAGAAGGCCCTTGAAGTCATCCCTGGAT	899
Db	943	GATGGAGAAGGACCCGAAATAAAGATACATCACGCTGTGAGCAGCAGCTCGGCACCCATGGAT	1002
Qy	900	TGACGGAAACAGGCCCTCCACCGGGACATCTACCATTCACTGTCAGCTCCAGATCCAGAA	959
Db	1003	CGCTGGTGACACAGCCCTACAACAAAACATCCACGAGTCCTGACGCGCCAGATCCGGAA	1062
Qy	960	GAACCTTTGCTTAAGACAAGTGGAGCAAGCCCTTAACGACGACGCTGTGGTGACACCAT	1019
Db	1063	AAACTTTGCCAAGAGCAAATGGAGACAAGCATTTTAATGCCACGGCGTCGTGAGACATAT	1122
Qy	1020	GAGGAAGCTACACAT	1034
Db	1123	GAGAAACTACACCT	1137
RESULT	12		
AXI	166517		
LOCUS	AXI166517	1074 bp	DNA linear PAT 22-JUN-2001
DEFINITION	Sequence 8 from Patent WO0138503.		
ACCESSION	AXI166517		
VERSION	AXI166517.1	GI:14546862	
KEYWORDS	.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 1074); Manning,G.S., Sudarsanam,S.S., Martinez,R.,		
JOURNAL	Flanagan,P. and Clardy,D.S.		
FEATURES	Novel human protein kinases and protein kinase-like enzymes		
source	Patent: WO 0138503-A 8 31-MAY-2001;		
BASE COUNT	Sugen, Inc. (US)		
ORIGIN	Location/Qualifiers		
	1..1074		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	313 a 258 c 276 g 227 t		
	Query Match 20.5%; Score 502.2; DB 6; Length 1074;		
	Best Local Similarity 71.3%; Pred. No. 1.4e-127;		
	Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;		
Qy	70	ATGGGTCAAAAGGARAAGATGACTGCGAGTTCCTGGRAAAACACACACCACCAATCCGG	129
Db	1	ATGGCCCCGGGAGACGGCGAGACAGCTTCCTCTGGAAAAAGCAAGCTGAAGACATCAAG	60
Qy	130	AAAACCTTCATTTTTATGGAAAGTGTGGGATCAGGAGCTTTCTCAGAAAGTTTTCTGTG	189
Db	61	AAGATCTTCGAGTTCAAAGAGACCCTCGAACCGGGGCTTTTCGGAAGTGGTTTTAGCT	120
Qy	190	AAGCAAAAGCTGACTGGGAAGCTCTTCGTCGAAGTGCATC---AAGAAGTCACCTGCC	246
Db	121	GAAAGAAAGGCAACTGGCAAGCTCTTCTGTGAAGTGTATCCCCTAAAGGCGGTGAAG	180
Qy	247	TTCCGGGACAGCAGCGCTCGGAATGAGATGTCTGTGTTGAAAAGATCAAGCATGAAAC	306
Db	181	GGCAAGHAACGACATAGAGAATGAGATAGCCGTCCTGAGAAAGATTAAAGCATGAAAT	240
Qy	307	ATTGTGACCCCTGGAGGACATCTATGAGAGCACACCACCTACTACTCTGGTCATGCACTT	366
Db	241	ATTGTTGCCCTGGAAAGCATTTATGAAAGCCCCAATACCTGTACTTGTGTCATGCAGCTG	300
Qy	367	GTTCCTGTGGGAGCTCTTTTGACCGGATCTGTGGAGCGGGGTGCTACACAGAGAAGAT	426
Db	301	GTGTCGGTGGNAGACTGTTTGACCGGATAGTGGAGAAGGGGTTTTATACAGAGAAGAT	360
Qy	427	GCCAGTCTGGTAGTCCAGAGGCTCTTTCGCGGAGTGAATACCTACATGAGAATGGCATC	486
Db	361	GCACGACTCTGATCCGGCAAGCTCTTGGACGCGGTGTACTACTCTCCACAGAATGGGCATC	420

QY 487 GTCACAGAGACTTAAGCCGCAAAACCTGCTTTACCTTACCCCTCAAGAGAACTCTAAG 546
DB 421 GTCCACAGAGACTCAAGCCGCAAAATCTCTTGACTACAGTCAAGATGAGGAGTCCAAA 480
QY 547 ATCATGATCACTGACTTTGGTCTGTGCTCAAGATGGAACAAATGG---CATCATGTCCACT 603
DB 481 ATAATGATCAGTACATTTGGATTGTCAAAAATGGAGGGCAAAAGAGATGTGATGTCCACT 540
QY 604 GCCTGTGGGACCCAGGCTACGTGGCTCCAGAGTCTGCGCCAGAAACCCCTACAGCAAG 663
DB 541 GCCTGTGGAATCCAGCTATGTCTGCTGAAGTCTGCGCCAGAAACCTTACAGCAAA 600
QY 664 CTGTGTGATGTCTGCTGCTCACTGCGGCTACATCATATATGCTGTGTGATGATCCGCCG 723
DB 601 GCGCTTGACTGCTGGTCCATCGGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 660
QY 724 TTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAGATCAAGAGGGCTTACTATGAGTTT 783
DB 661 TTTTATGATCAAAATGACTCCAAGCTTTTGAGCAGATCTCAAGGGGGAATATGAGTTT 720
QY 784 GAGTCTCCATTTCTGGGATGACATTTCTGAGTCAAGCAAGGACTTTATTTGCCACTTGTCT 843
DB 721 GACTCTCCCTTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCATTCGGAACCTGATG 780
QY 844 GAGAAGATCCGAACGAGCGGTACACTGTGAGAGGCGCTTGAGTCATCCCTGGATTGAC 903
DB 781 GAGAAGGACCGAATAAAGATACAGTGTGAGCAGAGCTCGGCAACCATGATGCTGCT 840
QY 904 GGAACACGCGCCCTCCACCGGGACATCTACCCATCAGTCAAGCTCCAGATCCAGAGAAC 963
DB 841 GGTGACACGCGCTCAACAAAACATCCAGAGTCCGTCAGCGCCAGATCCGGAHAAC 900
QY 964 TTTGCTTAAGAGCAAGTGGAGGCAAGCCTTCAAGCAGCAGCTGTGTGATGATGAGG 1023
DB 901 TTTGCCAAGAGCAATGGAGACAAAGCATTTAATGCCACGCGCTGCTGAGACATATGAGA 960
QY 1024 AAGCTACACAT 1034
DB 961 AACTACACCT 971

RESULT 13
AX167587
LOCUS AX167587 1074 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 3 from Patent WO0142435.
ACCESSION AX167587
VERSION AX167587.1 GI:14596987
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1074)
AUTHORS Donoho, G., Scoville, J., Turner, C.A., Friedrich, G., Zambrowicz, B., Abuin, A., and Sands, A.T.
TITLE Novel human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0142435-A 3 14-JUN-2001;
LEXON Lexicon Genetics Incorporated (US)
FEATURES
source
location/Qualifiers
1..1074
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 313 a 258 c 276 g 227 t
ORIGIN

Query Match 20.5%; Score 502.2; DB 6; Length 1074;
Best Local Similarity 71.3%; Pred. No. 1.4e-127;
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;

QY 70 ATGGGTCGAAGAAGATGATGACTGCTGCTGGAAGAACAGACCAACCAATCCGG 129
DB 1 ATGGCCGGGAGACGCGGAGAGCAGCTCTCTCTGGAAAAGCAAGCTGAAGACATCAAG 60

QY 130 AAAACCTTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTTTTCCTGGTG 189
DB 61 AAGATCTTCAGTTCAAAAGAGACCCTCGGAACCGGGCCCTTTCCGAAGTGGTTAGCT 120
QY 190 AAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAAGTGCATC---AAGAAGTCACCTGCC 246
DB 121 GAAGAGAAGCAACTGCAAGCTCTTTGCTGTGAAGTGTATCCCTAAGAAGCGCTGAAG 180
QY 247 TTCGGGACAGCGCTGGAGAATGAGATGTCTGTGTTGAAAAGATCAAGCATGAAGAAC 306
DB 181 GGCAAGGAAGCAGCAGATAGAGATGAGATGAGTCCCTGAGAAGATTAAAGCATGAAT 240
QY 307 ATTGTGACCTGGAGGACATCTATGAGAGCACCACCCTACTACTGTGTGATGAGCTTT 366
DB 241 ATTGTTCCTGGAAGACATTTATGAAGCCCAATCAGCTGTACTTGTGTGATGAGCTGTG 300
QY 367 GTTCTCTGCTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGAT 426
DB 301 GTGTCGCTGGAGAGCTTTTGACCGGATAGTGAAGAGGGTGTATACAGAGAAGGAT 360
QY 427 GCCAGTCTGTGTGATGCCAGAGTCTTGTGCGCAGTGAATATACATGAGATGGCATC 486
DB 361 GCCAGCACTCTGATCCGCCAAGTCTTGGAGCGCTGTACTATCTCCACAGAATGGGCATC 420
QY 487 GTCCACAGAGACTTAAAGCCGGAACCTGCTTTACCTTACCCTCAAGAGACTCTAAG 546
DB 421 GTCCACAGAGACTTCAAGCCCGAAATCTCTGTACTACAGTCAAGATGAGGATCCAAA 480
QY 547 ATCATGATCACTGACTTTGCTCTCAAGATGGAACAGAAATGG---CATCATGTCCACT 603
DB 481 ATATGATCAGTGAATTTGATTTGCAAAATGGAGGCAAGAGATGTGATGCTCACT 540
QY 604 GCCTGTGGGACCCAGGCTACGTGGCTCCAGAGTGTGCGCCAGAAACCCCTACAGCAAG 663
DB 541 GCCTGTGGAATCTCAGGCTATGCTGCTTGAAGTCTCGCCAGAAACCTTACAGCAAA 600
QY 664 GCTGTGATGCTGGTCCATCGGCTCATCAGCTACATATTTGCTGTGTGATACCCCGG 723
DB 601 GCGCTTGAAGTCTGCTGCTCAAGATGGAACAGAAATGG---CATCATGTCCACT 660
QY 724 TTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTTACTATGAGTTT 783
DB 661 TTTTATGATGAATGACTCCAAAGCTCTTTGAGCAGATCTCAAGGCGGAATATGAGTTT 720
QY 784 GAGTCTCCATTTCTGGGATGACATTTCTGAGTCAAGCAAGGACTTTATTTGCCACTTGT 843
DB 721 GACTCTCCCTTACTGGGATGACATCTCGGACTCTGCAAAAGACTTCATTCGGAACCTGATG 780
QY 844 GAGAAGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTTGAGTCACTCCCTGGATTGAC 903
DB 781 GAGAAGGACCCGAATAAAGATACAGCTGTGAGCAGCAGCTCGGCAACCATGATCGCT 840
QY 904 GGAAGACGCGCTCCACCGGGACATCTACCATCAGTCAAGCTCCAGATCCAGAGAAC 963
DB 841 GGTGACACGCGCTCAACAAAACATCCAGAGTCCGTCAGCGCCAGATCCGGAHAAC 900
QY 964 TTTGCTTAAGAGCAAGTGGAGGCAAGCCTTCAAGCAGCAGCTGTGTGATGATGAGG 1023
DB 901 TTTGCCAAGAGCAATGGAGACAAAGCATTTAATGCCACGCGCTGCTGAGACATATGAGA 960
QY 1024 AAGCTACACAT 1034
DB 961 AACTACACCT 971

RESULT 14
AX167585
LOCUS AX167585 1158 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 1 from Patent WO0142435.
ACCESSION AX167585
VERSION AX167585.1 GI:14596986
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1158)
Doehho.G., Scoville.J., Turner.C.A., Friedrich.G., Zambrowicz.B., Abuin.A. and Sands.A.P.
TITLE Novel human kinase proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0142435-A 1 14-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
1..1158
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 323 a 283 c 304 g 248 t
ORIGIN
Query Match 20.5%; Score 502.2; DB 6; Length 1158;
Best Local Similarity 71.3%; Pred. No. 1.4e-127;
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;
Qy 70 ATGGGTGGAAGAGAGATGACTGTCAGTTCCTGGAAGAAACAGACCAACATCCGG 129
Db 1 ATGGCCCGGAGACGGGAGAGCAGCTCTCTCTGGAAGAAAGAGCTGAAGACATCAAG 60
Qy 130 AAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCAGAGCTTTCTCGTGTG 189
Db 61 AAGATCTTCGAGTTCAAGAGAGCCCTCGGAACCGGGGCTTTTCCGAGAGTGGTTTAGCT 120
Qy 190 AAGCAAGACTGACTGGGAAGCTTTTGCTCTGAAGTGCATC---AAGAAGTCACTCGCC 246
Db 121 GAAGAGAGGCAACTGGCAAGCTTTTGCTGTGAAGTGTATCCCTAAGAGGCGCTGAAG 180
Qy 247 TTCGGGACACGCTGGAGAAATGAGATGCTGTGTTGAAAAGATCAGCATGAAGAAC 306
Db 181 GGCAAGGAAGCAGCATAGAAATGAGATAGCCGCTCTGAGAAGATTAAGCATGAAGAT 240
Qy 307 ATTGTGACCTGGAGGACATCTATGAGAGCACCAACCTACTACCTGGTGCATGCAGCTT 366
Db 241 ATTGTGCTGGAGACATTTATGAAGCCCAATCACCTGTACTTGGTGCATGCAGCTG 300
Qy 367 GTTCTGTGGGGAGCTCTTTGACCGGATCTGGAGCGGGGTGCTACAGAGAGGAT 426
Db 301 GTGTCGGGTGAGAGCTGTTTGACCGGATAGTGAGAGGGGGTTTATACAGAGAGGAT 360
Qy 427 GCCAGCTGTGATCCAGCAGCTCTGTGCGCAGTGAATACCTACATGAGATGGCATC 486
Db 361 GCCAGACTGTATCCGCCAAGCTTTGGACCGCTGTACTATCTCCACAGATGGGCATC 420
Qy 487 GTCCACAGAGACTTAAAGCCGAAACCTGCTTTTACCTTACCCCTGAAGAGACTCTAAG 546
Db 421 GTCCACAGAGACTCAAGCCGAAATCTCTTGTACTACAGTCAAGATGAGGATCCAAA 480
Qy 547 ATCATGATCACTGACTTTGGTCTGTCAAGATGGAACAGAAATGG---CATCATGTCCACT 603
Db 481 ATAATGATCAGTACTTTGGATTGTCAAAAATGAGGCGCAAGAGAGATGTGATGCCACT 540
Qy 604 GCCTGTGGAGCCAGGCTACGCTCCAGAGTGTGCGCCAGAACCTACAGCAAG 663
Db 541 GCCTGTGGAATCCAGGCTATGTGCTCTCTGAACTCTCGCCCAAGAACTTACAGCAA 600
Qy 664 GCTGTGATTTGCTGCTCCATCGGGCTCATCACCTACATATTTGCTGTGGATACCCCGG 723
Db 601 GCGGTTGACTGCTGCTCATCGGAGTGATTCCTTACATCTTGCTCTCGGGCTACCTCCT 660
Qy 724 TTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAGTTT 783
Db 661 TTTTATGATGAATGATCTCAAGCTCTTTGAGCAGATCTCTCAAGGCGGAATATGAGTTT 720
Qy 784 GAGTCTCCTACTTGGGATGACATTTCTGAGTCAGCCAGGACTTTATTTGCACTTCTCTT 843
Db 721 GACTCTCCTACTGGGATGACATCTCCGACTCTCCAAAGACTTCATTCGGAACCTGATG 780

Qy 844 GAGAAGATCCGAACGAGCGGTACACCTGTGAGAAGGCCTTGAGTCATCCCTGGATTGAC 903
Db 781 GAGAAGAGCCCGAATAAATAATACACGTGTGACGACAGCTCGGCACCCATGGATCGCT 840
Qy 904 GGAACACAGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAAGAAC 963
Db 841 GGTGACACAGCCCTCAACAAAACATCCACGAGTCGCTCAGGCCAGATCCGGAAGAAC 900
Qy 964 TTTGCTTAAGAGCAAGTGGAGCAAGCCTTCAACGACGAGCTGTGTGTACCAATGAGG 1023
Db 901 TTTGCCAAGAGCAATGAGAGCAAGCATTTAATGCCACGCGCTGCTGAGACATATGAGA 960
Qy 1024 AAGCTACACAT 1034
Db 961 AACTACACCT 971
RATCAMPKAA 1439 bp mRNA linear ROD 18-MAY-1995
Rattus norvegicus Cam-like protein kinase mRNA, complete cds.
L26288
VERSION L26288.1 GI:439613
KEYWORDS protein kinase.
SOURCE Rattus norvegicus (strain Sprague-Dawley) (library: day 21
gestation lung cDNA) fetal lung cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1439)
AUTHORS Cho.F.S., Phillips.K.S., Bogucki.B. and Weaver.T.E.
TITLE Characterization of a rat cDNA clone encoding
calcium/calmodulin-dependent protein kinase I
JOURNAL Biochim. Biophys. Acta 1224 (1), 156-160 (1994)
MEDLINE 95035115
PUBMED 7948038
FEATURES
Location/Qualifiers
1..1439
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="lung"
/dev_stage="fetal"
/tissue_lib="day 21 gestation lung cDNA"
118..1242
/codon_start=1
/product="Cam-like protein kinase"
/protein_id="AAA66944.1"
/translation="MPGAVEGPRNKAQEDIRIDYFDVLGTGAFSEVILAEKRTQK
LVAIKCIAKALEGEGSMENFIAVLHKIKNPVALDDIYESGSHLYLIMQLVSGGE
LFDRIVEKGYTERDASRLIFQVLDVAVYLDLGIYHRDLKPNLLYYSLDEDSKIMI
SDFGUSKMEDEPGLSVLTACGTPGYVAPEVLAQKPYSAKDFIRHLMKDKPKRFTCEALQHPWI
YDENDAKLIFEQILKAEYBFDSPDDISDAKDFIRHLMKDKPKRFTCEALQHPWI
AGDTALDKNIHQVSEQIKKNFAKSKOAFNATAVVRMRKQLQGTSGEQGGGTASH
GELLPTAGGPAAGCCCRDCCVPEPSKELPPPPSSRAM"
BASE COUNT 357 a 375 c 405 g 302 t
ORIGIN
Query Match 20.1%; Score 493; DB 10; Length 1439;
Best Local Similarity 71.7%; Pred. No. 5.2e-125;
Matches 675; Conservative 0; Mismatches 260; Indels 6; Gaps 2;
Qy 101 CTGTGAAGAAACAGACACCACACATCCGGAACACCTTCATTTTATGGAAGTCTGGGAT 160
Db 140 CCAGGTGGAAGCAGGAGAGACATAGGGATATTTATGACTTCAGGATGTTCTGGGCA 199
Qy 161 CAGGAGCTTCTCAGAAAGTTTCTCGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTC 220
Db 200 CGSGTGCCTTCTCAGAAAGTATCTCTGGCAGAGGACAGAGGACTCAGAAACTGGTGCCA 259
Qy 221 TGAAGTGCAT---CAAGAAGTCACTGCTTCCTCCGGACAGACGCTGGAGAATGAGATTG 277

Search completed: March 14, 2003, 17:25:23
Job time : 7400 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 17:36:01 ; Search time 328 Seconds
(without alignments)
3268.144 Million cell updates/sec

Title: US-09-960-643-2

Perfect score: 2513
Sequence: 1 MGRKEEDCSNKKQTTNIR.....VKASGSSHCRAGOTGVCLIM 476

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09960643/runat_07032003_090517.19867/app_query.fasta_1.647
-DB=N_Geneseq_101002 -OFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09960643 -CGN_1_1_263 -runat_07032003_090517.19867 -NCFU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	2513	100.0	1956	22	AAI60703	Human polynucleoti
2	2513	100.0	2447	24	AAD36140	Human calmodulin k
3	2499.5	99.5	2165	22	AAI58917	Human polynucleoti
4	2327.5	92.6	2689	22	AAS31014	Human diagnostic a
5	1250.5	49.8	1736	22	AAD18817	Human kinase (PRIN
6	1246	49.6	1074	22	AAS06708	Polynucleotide seq
7	1246	49.6	1074	22	AAH25119	Nucleotide sequenc
8	1246	49.6	1158	22	AAH25118	Nucleotide sequenc
9	1246	49.6	1565	24	ABL60905	Human CAM kinase I
10	1246	49.6	1661	22	ABL60905	Human kinase (PRIN
11	1246	49.6	1671	22	AAH25120	Nucleotide sequenc
12	1246	49.6	2164	22	AAI60424	Human polynucleoti
13	1243	49.5	1733	22	AAI58638	Human polynucleoti
14	1238.5	49.3	1694	22	AAC90432	Murine Dendritic C
15	1237	49.2	1578	22	AAD04775	Human death domain
16	1150.5	45.8	1728	21	AZ46162	cDNA sequence enco
17	1070	42.6	1458	19	AAV18867	Calmodulin-depende
18	1041	41.4	1032	22	AA06711	Polynucleotide seq
19	1036	41.2	1512	23	AA070157	DNA encoding novel
20	990	39.4	1282	20	AA06835	Disease associated
21	957.5	38.1	2264	23	ABL02719	Drosophila melanog
22	753	30.0	2494	22	AAK94614	Human full-length
23	753	30.0	2747	22	AAI58143	Human polynucleoti
24	753	30.0	2897	22	AAF30481	Human protein phos
25	753	30.0	2905	22	AAK94325	Human full-length
26	747.5	29.7	2297	22	AAH78261	Nucleotide sequenc
27	747.5	29.7	3017	22	AAD18832	Human kinase (PRIN
28	747.5	29.7	3124	24	ABA95682	Human protein kina
29	747	29.7	2840	22	AAI58144	Human polynucleoti
30	733.5	29.2	3022	22	AAK94299	Human full-length
31	726.5	28.9	1572	22	AAH99632	Human protein enco
32	683.5	27.2	1774	23	AA068767	DNA encoding novel
33	683.5	27.2	1774	23	AA090919	DNA encoding novel
34	667	26.5	1763	23	AA066250	DNA encoding novel
35	664.5	26.3	3471	18	AAAT71761	PSKH-1 cDNA. Homo
36	661.5	26.3	1764	23	AA064534	DNA encoding novel
37	650.5	25.9	1386	22	AAK52586	Human polynucleoti
38	650.5	25.9	1450	22	AAK51602	Human polynucleoti
39	639.5	25.4	3552	24	ABL40502	Human serine/threo
40	633.5	25.2	2667	24	AAD34311	Human PKIN-14 cDNA
41	625.5	24.9	2301	24	ABL40503	Human serine/threo
42	615.5	24.5	2110	21	AAZ46160	cDNA sequence enco
43	611.5	24.3	1803	22	AAH46903	cDNA encoding huma
44	611.5	24.3	1803	24	AAD38861	Human kinase (PRIN
45	611.5	24.3	7626	24	AET04063	Human ovary specif

ALIGNMENTS

RESULT 1

AAI60703
ID AAI60703 standard; cDNA; 1956 BP.
XX AAI60703;
AC AAI60703;
XX
XX
XX 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 4692.

XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW ankyrotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.

XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX P-PSDB; AAM41547.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 4692; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 1956 BP; 474 A; 559 C; 508 G; 414 T; 1 other;

Alignment Scores:
Pred. No.: 5,24e-206 Length: 1956
Score: 2513.00 Matches: 476
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-960-643-2 (1-476) x AAI60703 (1-1956)

QY 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrThrAsnIleArg 20
DB 65 ATGGGTGGAAGGAGCAAGATGACTGCTCTCTGGAAGAAACACACCACCAATCCGG 124
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
DB 125 AAAACCTTCATTTTATGGAAGTGTCTGGGATCAGGAGCTTCTCAGAAAGTTTCTCGTGTG 184
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
DB 185 AAGCAAGAGTACTGGGAGGCTCTTTGCTCTGAAAGTGCATCAAGAAAGTCACTGCCCTTC 244
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80

DB 245 CGGCACAGCAGCCTGGAGAAATGAGATTGCTGTGTAAGAAAGATCAAGCATGAAAAACATT 304
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
DB 305 GTGACCTCTGGAGGACATCTATGAGAGCACCACTACTACTCTGCTGCTACGAGCTTGT 364
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
DB 365 TCTGTTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATGCC 424
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
DB 425 AGTCTGTGTATCCAGCAGCTTGTGCGCAGTGAATACCTACATGAGATGATCGCTGC 484
QY 141 HisArgAspLeuLysProGluAsnLeuTyrLeuThrProGluGluAsnSerLysIle 160
DB 485 CACAGAGACTTAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAGTCTAAGATC 544
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
DB 545 ATGATCACTGACTTTGTGCTGTCCAGATGGAACAGAAATGGCATCATGTCCACTGCTGT 604
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
DB 605 GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCTACACAGAGGCTGTG 664
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220
DB 665 GATTGCTGGTCCATCGGCGTCATCACCTACATATTGCTGTGTGATACCCGCCCATCTAT 724
QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
DB 725 GAAGAAACGGAGTCTAAGCTTTTCGAGAGATCAAGAGGGGCTACTATGATGTTGAGTCT 784
QY 241 ProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
DB 785 CCATCTCGGATGACATTCTGAGTCAGCCAGGACTTTATTGCCACTTGTGAGAAG 844
QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
DB 845 GATCCGAACGAGCGGTACACTGTGTGAGAAGCCCTTGAGTCACTCCCTGGATTGACGAAAC 904
QY 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
DB 905 ACAGCCCTCCACCGGGACATCTACCCATCAGTCAGCTCCAGATCCAGAGAAGACTTGTCT 964
QY 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320
DB 965 AAGAGCAAGTGGAGGCAAGCCTTCAACGCAGCAGCTGTGTCACCAACATGAGGAAGCTA 1024
QY 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr 340
DB 1025 CACATGAACCTGACAGCGCGGCGTCCGCCAGAGGTGGAGAACAGCGCGCTGAAACT 1084
QY 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360
DB 1085 CAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCAGGAGGACCTGTC 1144
QY 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
DB 1145 CTGACCAACAGTGTAGCAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1204
QY 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
DB 1205 CCCACTGCCCTGTGTGCGAGTCCCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1264
QY 401 SerSerLeuValProMetHisGlnLysSerLeuAlaAlaGlyProCysGlyCysSer 420
DB 1265 AGCAGCCTGTGTGCCCATGCATCAGGGGTCCCTGCGCCCGCGGCGCTGTGCTGCTGCT 1324
QY 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu 440
DB 1325 AGCTGCTGAACTTGGGAGCAAGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1384

QY 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460
AAD36140
ID AAD36140 standard; DNA; 2447 BP.
XX
AC AAD36140;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human calmodulin kinase, CAMK-X1 gene.
XX
KW Human; cytostatic; antisense gene therapy; screening; protein kinase;
KW cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-X1;
KW calmodulin kinase; enzyme; gene; chromosome lq32.1-32.3; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 70..1500
FT /*tag= a
FT /product= "Human CAMK-X1 protein"
FT
XX WO200224947-A2.
XX
XX 28-MAR-2002.
XX
XX 20-SEP-2001; 2001WO-IB02237.
XX
XX 20-SEP-2000; 2000US-233999P.
XX 02-OCT-2000; 2000US-237419P.
XX 02-OCT-2000; 2000US-237423P.
XX 04-OCT-2000; 2000US-238558P.
XX 10-MAY-2001; 2001US-290555P.
XX (KINE-) KINETEK PHARM INC.
PA (UYBR-) UNIV BRITISH COLUMBIA.
XX
XX Yoganathan T, Delaney AD;
XX
DR WPI; 2002-394145/42.
XX P-PSDB; AAE22764.
XX
XX Diagnosing cancer, comprises determining the upregulation of expression
PT of a nucleic acid sequence encoding a protein kinase or upregulation of
PT expression of the protein kinase, in the cancer
XX
PS Claim 16; Page 62-64; 87pp; English.
XX
XX The invention relates to a method for screening biologically active agent
CC that modulates cancer associated protein kinase function. The invention
CC also relates to a method for diagnosing cancer comprising determining the
CC upregulation of expression of a nucleic acid sequence encoding a protein
CC kinase. The method is useful for diagnosing cancer. A protein kinase is
CC useful for screening biological agents that modulate cancer associated
CC protein kinase function. Downregulating the activity of protein kinase is
CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon
CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy
CC derived tumours and inflammatory samples such as arthritic synovium, for
CC amplified DNA in the cell or increased expression of corresponding mRNA
CC or protein and is also useful to detect differences in expression levels
CC such as molecular weight, amino acid and nucleotide sequences between the
CC two cells. The present sequence is human calmodulin kinase CAMK-X1 gene
CC located on chromosome lq32.1-32.3.
XX
XX Sequence 2447 BP; 590 A; 707 C; 604 G; 546 T; 0 other;

Alignment Scores:
Pred. No.: 7.2e-206 Length: 2447
Score: 2513.00 Matches: 476
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
US-09-960-643-2 (1-476) x AAD36140 (1-2447)
QY 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 70 ATGGTCGAAAGGAAGAGATGACTGCAGTTCTCGGAAGAAACAGACACCATCCGG 129
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 130 AAAACCTTCATTTTATGGAAGTCTGGATCAGAGCTTCTCAGAAGTTTCTCCTGGTG 189
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 190 AAGCAAGACACCTGACTGGGAGCTCTTTGCTCTGAAGTGCATCAAGAAGTACCTGCCTTC 249
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
Db 250 CGGACACAGCAGCCGGAGATGAGATTGCTGTGTGAAAGATCAAGACATGAACATT 309
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 310 GTGACCTGGAGGACATCTATGAGACAGCACACCACCTACTACCTGGTGCATGCAGCTTGT 369
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Db 370 TCTGGTGGGAGCTCTTTGACCGGATCTGGAGCGGGTGTCTACACAGAGAAGGATGCC 429
QY 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 430 AGTCTGGTGATCCAGCAGGTCTTCTCGCAGTGAATACCTACATGAGATGGCATCGTC 489
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 490 CACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTAAGATC 549
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 550 ATGATCACTGACTTGTGTCTCCAGATGAACAGATGGCATCATGTCCACTGCCCTGT 609
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 610 GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCCCAGAAACCCCTACAGCAAGGCTGTG 669
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220
Db 670 GATTGCTGGTCCATCGGCGTCATCACCCTACATATGTCTGTGGATACCCCGCTTCAT 729
QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Db 730 GAACAAAGGAGTCAAGCTTTTCGACAGATCAAGGAGGCTACTATGATTTGAGTCT 789
QY 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
Db 790 CCATTCTGGGATGACATTTCTGAGTCAGCCAAAGGACTTTATTGCCACTTGTCTGGAAG 849
QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
Db 850 ATCCGACAGCGCGGTACCTGTGAGAAGCCCTTGAGTTCATCCCTGGATTGACGGAAC 909
QY 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
Db 910 ACGGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAGAACTTTGCT 969
QY 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320
|||||

Db 970 AGAGCAAGTGGAGGCAAGCCCTTCAACGACAGAGCTGTGGTGCACACCATGAGGAAGCTA 1029

QY 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProGluThr 340

Db 1030 CACATGAACCTGCACAGCCGGCGCTCCGCCAGAGGTGGAGAACAGCGCGCTGAACCT 1089

QY 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360

Db 1090 CAAGCCTCAGAAACCTCTAGACCAGCTCCCTGAGATCACCATCAGCGAGGACCTGTC 1149

QY 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380

Db 1150 CTGGACACAGGTAGACATCCCTGCCCTGACCCCAATTACCTGGCAGATGGCGCGGG 1209

QY 381 ProThrAlaProGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400

Db 1210 CCCACTGCCCTGGTGGCAGGTCCCTCACTGCTGGTCAATGGCTCCCTCCACATCAGC 1269

QY 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420

Db 1270 AGCAGCTGTGCCATGCATCAGGGGTCCCTGGCGCGCGCGCTGTGGCTGCTGCTCC 1329

QY 421 SerCysLeuAsnIleGlySerGlyLysSerSerTyrCysSerGluProThrLeuLeu 440

Db 1330 AGCTGCTGAACATTTGGGAGCAAGAAAGTCTCTACTGCTCTGAGCCACACTCCTC 1389

QY 441 LysLysAlaAsnLysLysGlnAsnPhelLysSerGluValMetValProValLysAlaSer 460

Db 1390 AAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTATGTACCATGTTAAAGCCAGT 1449

QY 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476

Db 1450 GGCAGCTCCCACTGCCGGCAGGCGCAGACTGGAGTCTGTCTCATTTATG 1497

RESULT 3

ID AAI58917 standard; cDNA; 2165 BP.

XX AAI58917;

XX 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 1120.

DE Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

XX leukaemia; ss.

XX Homo sapiens.

OS WO200153312-A1.

PN 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPL; 2001-442253/47.

DR P-PSDB; AAM39761.

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX Claim 1; SEQ ID NO 1120; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX SQ Sequence 2165 BP; 533 A; 613 C; 563 G; 455 T; 1 other;

Alignment Scores:

Pred. No.: 8.74e-205 Length: 2165

Score: 2499.50 Matches: 476

Percent Similarity: 98.55% Conservative: 0

Best Local Similarity: 98.55% Mismatches: 0

Query Match: 99.46% Indels: 7

DB: 22 Gaps: 1

US-09-960-643-2 (1-476) x AAI58917 (1-2165)

QY 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20

Db 253 ATGGTCGAAAGGAAGAGATGACTGCAGATTCTCTGGAAGAACAGACCACCAATCCGG 312

QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40

Db 313 AAAACCTTCATTTTATGGAAGTGTGGGATCAGAGCTTTCTCAGAAGTTTCTCGTGGT 372

QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60

Db 373 AAGCAAAAGCTGACTGGGAAGCTCTTTCCTCTGAAGTGCATCAAGAGTCACCTGCCTTC 432

QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80

Db 433 CGGGACAGCAGCTGGAGAATGAGATTGCTGTGTGAAAAAGATCAAGCATGAAACATT 492

QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100

Db 493 GTGACCTGGAGGACATCTATGAGAGACCAACCCACTACTACCTGGTGCATGCACGTTGTT 552

QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120

Db 553 TCTGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAGAAGATGCC 612

QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140

Db 613 AGTCTGGTGATCCAGCAGGTCTTTGTGGCAGTGAATAACCTACATGAGAATGGCATCGTC 672

QY 141 HisArgAspLeuLys-----ProGluAsnLeuLeuTyrLeuThr 153

Db 673 CACAGAGACTTAAGGTTCAAGCGGGAGTCTCGGCCCGGAAACCTCTTACCTTACC 732

QY 154 ProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsn 173

Db 733 CCTGAAGAGAACTCTAAGATCATGATCATCTACTGCTTTGGTGTCTCTCCAGATGAACAGAA 792

```
Qy 174 GlyIleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValIleuAlaGln 193
Db 793 GGCATCATGTCCTACTGCTGGACCCCGAGATAGTGCTGCCAGAGTGTGCCCGCAG 852
Qy 194 LysProTyrSerLysAlaValAspCysTyrSpSerIleGlyValIleThrTyrIleLeuLeu 213
Db 853 AAACCCCTACAGCAAGCTGTGGATTGCTGCTCATCGCGCTCATCACCTACATATTGCTC 912
Qy 214 CysGlyTyrProProPheTyrGluThrGluSerLysLeuPheGluLysIleLysGlu 233
Db 913 TGTGGATACCCCTCTATCAAGAAACGGAGCTTAAGCTTTTCGAGAAGATCAAGGAG 972
Qy 234 GlyTyrTyrGluPheGluSerProPheTyrAspAspIleSerGluSerAlaLysAspPhe 253
Db 973 GGCTACTATGAGTTTGGATCTCCATTCTGGATGACATTCTTGAGTCAGCCAAAGGACTT 1032
Qy 254 IleCysHisLeuLeuGluLysAspProAsnGluAtrGlyTyrThrCysGluLysAlaLeuSer 273
Db 1033 ATTTGCCACTTGTGAGAGGATCGGAACGAGCGGTACACCTGTGAGAAGGCCCTTGAGT 1092
Qy 274 HisProTyrIleAspGlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeu 293
Db 1093 CATCCCTGGATTACGGAACACAGCCCTCCACCGGGACATCTACCCATCAGTCAGCCTC 1152
Qy 294 GlnIleGlnLysAsnPheAlaLysSerLysTyrArgGlnAlaPheAsnAlaAlaVal 313
Db 1153 CAGATCCAGAAAGAACTTTGCTAGAGCAAGTGGAGCAAGCCTTCAACGCGCAGCAGCTGT 1212
Qy 314 ValHisHisMetArgLysLeuHisMetAsnLeuHisSerProGlyValArgProGluVal 333
Db 1212 GTGCACCATATGAGGAGCTACATGAACTGACAGCCCGGGCGTCCGCCACAGGTG 1272
Qy 334 GluAsnArgProProGluThrGlnAlaSerGluThrSerArgProSerSerProGluIle 353
Db 1273 GAGAACAGCGCGCTGAACTCAAGCCTCAGAACCTCTAGACCCAGCTCCCTCAGATC 1332
Qy 354 ThrIleThrGluAlaProValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLeu 373
Db 1333 ACCATCACCGAGGACCTGTCTCGACACAGTGTAGCAGCTCCCTGCGCTGACCCAAATTA 1392
Qy 374 ProCysGlnHisGlyArgProThrAlaProGlyGlyArgSerLysAsnCysLeuVal 393
Db 1393 CCTGCGCAGCATGGCGCGGCCACTGCGCTGTGTGGCAGGTCCCTCAACTGCCTGTGTC 1452
Qy 394 AsnGlySerLeuHisIleSerSerLeuValProMetHisGlnGlySerLeuAlaAla 413
Db 1453 AATGGCTCCCTCCACATCAGCAGCAGCTGGTGCCCATCATCGGGGTCCCTGGCGCGCC 1512
Qy 414 GlyProCysGlyCysSerSerCysLeuAsnIleGlySerLysGlyLysSerTyr 433
Db 1513 GGGCCCTGTGCTGCTGCTCCAGCTGCCTGAACATTTGGAGCAAGGAAGTCCCTCTAC 1572
Qy 434 CysSerGluProThrLeuLeuLysLysAlaAsnLysLysGlnAsnPhelLysSerGluVal 453
Db 1573 TGCTCTGAGCCCACTCTCTCAAAAAGGGCAACAAAAACAGAACTTCAAGTCGGAGGTC 1632
Qy 454 MetValProValLysAlaSerGlySerSerHisCysArgAlaGlyGlnThrGlyValCys 473
Db 1633 ATGTACAGTTAAAGCAGTGGCAGCTCCACTGCGGGCAGGCGAGACTGGAGTCTGT 1692
Qy 474 LeuIleMet 476
Db 1693 CTCATTATG 1701
RESULT 4
AAS31014
ID AAS31014 standard; cDNA; 2689 BP.
XX AC AAS31014;
XX XX
XX 04-DEC-2001 (first entry)
XX
```

```
DE XX
XX Human diagnostic and therapeutic polynucleotide (DITHP) #29.
KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;
KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;
KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;
KW respiratory disorder; ss.
XX OS Homo sapiens.
XX WO200162927-A2.
XX 30-AUG-2001.
XX 21-FEB-2001; 2001WO-US06059.
XX 24-FEB-2000; 2000US-0184693.
XX 24-FEB-2000; 2000US-0184697.
XX 24-FEB-2000; 2000US-0184698.
XX 24-FEB-2000; 2000US-0184768.
XX 24-FEB-2000; 2000US-0184769.
XX 24-FEB-2000; 2000US-0184770.
XX 24-FEB-2000; 2000US-0184771.
XX 24-FEB-2000; 2000US-0184772.
XX 24-FEB-2000; 2000US-0184773.
XX 24-FEB-2000; 2000US-0184774.
XX 24-FEB-2000; 2000US-0184776.
XX 24-FEB-2000; 2000US-0184777.
XX 24-FEB-2000; 2000US-0184797.
XX 24-FEB-2000; 2000US-0184813.
XX 24-FEB-2000; 2000US-0184837.
XX 24-FEB-2000; 2000US-0184841.
XX 24-FEB-2000; 2000US-0185213.
XX 24-FEB-2000; 2000US-0185216.
XX 12-MAY-2000; 2000US-0203785.
XX 15-MAY-2000; 2000US-0204226.
XX 16-MAY-2000; 2000US-0204525.
XX 16-MAY-2000; 2000US-0204821.
XX 16-MAY-2000; 2000US-0204908.
XX 16-MAY-2000; 2000US-0205232.
XX 17-MAY-2000; 2000US-0204815.
XX 17-MAY-2000; 2000US-0204863.
XX 17-MAY-2000; 2000US-0205221.
XX 17-MAY-2000; 2000US-0205285.
XX 17-MAY-2000; 2000US-0205286.
XX 17-MAY-2000; 2000US-0205287.
XX 17-MAY-2000; 2000US-0205323.
XX 17-MAY-2000; 2000US-0205324.
XX (INCY-) INCYTE GENOMICS INC.
XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
XX Chen A, D'Sa SA, Anshey S, Dahl CR, Dam TC, Daniels SE;
XX Dufour GE, Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL;
XX Liu TF, Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;
XX Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;
XX Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;
XX WPI: 2001-502867/55.
XX P-PSDB: AAU19443.
XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.
XX enzymes, hormones and receptors, useful in diagnostics and therapeutics
XX
XX Claim 1; Page 310-311; 522pp; English.
XX The invention relates to polynucleotides (I) encoding diagnostic and
XX therapeutic (DITHP) polypeptides (II), which include e.g. enzymes,
XX and proteins involved in growth and development and receptors. (I) and
XX (II) may be used in the prevention, diagnosis and treatment of diseases
XX associated with inappropriate DITHP expression. For example, (I) and
XX (II) may be used to treat disorders associated with decreased polypeptide
XX expression by rectifying mutations or deletions in a patient's genome,
```


CC that affect the activity of the DITHPs, by expressing inactive proteins
 CC or supplementing the patient's own production of them. (I) and (II)
 CC may be used to treat diseases, for example, cell proliferative disorder,
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,
 CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. (I) and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. (II) may also be used as antigens in the production of
 CC antibodies against DITHPs and in assays to identify modulators of DITHP
 CC expression and activity. The anti-DITHP antibodies and antagonists may
 CC also be used to down regulate expression and activity. The anti-DITHP
 CC antibodies may also be used as diagnostic agents for detecting the
 CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbant
 CC assay (ELISA)). AAS30986-AAS31196 represent human diagnostic and
 CC therapeutic (DITHP) polynucleotides of the invention.

XX SQ Sequence 2689 BP; 642 A; 748 C; 678 G; 621 T; 0 other;

Alignment Scores:

Pred. No.: 7,06e-190 Length: 2689
 Score: 2327.50 Matches: 460
 Percent Similarity: 95.86% Conservative: 3
 Best Local Similarity: 95.24% Mismatches: 13
 Query Match: 92.62% Indels: 7
 DB: 22 Gaps: 1

US-09-960-643-2 (1-476) x AAS31014 (1-2689)

QY 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrThrAsnIleArg 20
 DB 237 ATGGTTCGAAGGAGAGAGTGAAGTCACTTCTTGGAGAAACAGACACCAACATCCGG 296
 QY 21 LysThrPheIlePheKetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
 DB 297 AAAACCTTCATTTTATGGAAGTGCCTGGGATCAGGAGCTTCTCAAGAGTTTTCCTGGTG 356
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
 DB 357 AAGCAAGACTGACTGGGAGCTCTTGGCTCTGAAGTGCATCAGAAGTCACCTGCCCTTC 416
 QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
 DB 417 CGGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAAACATT 476
 QY 81 ValThrLeuGluAspIleTrpGluSerThrThrHisTrpLysValMetGlnLeuVal 100
 DB 477 GTGACCTGGAGGACATCTATGAGACACACCCCTACTACTGTCTATGAGCTTGTGT 536
 QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTrpThrGluLysAspAla 120
 DB 537 TCTGGTGGGAGCTCTTGAACCGATCTCTGGAGCGGGTGTCTACACAGAGAAGATGCC 596
 QY 121 SerLeuValIleGlnValLeuSerAlaValLysTrpLeuHisGluAsnGlyIleVal 140
 DB 597 AGTCTGGTGATCCAGCAGGTCTTGTCCGAGATGAAATACATACAGAAATGGCATCGTC 656
 QY 141 HisArgAspLeuLysProGluAsnLeuLeuTrpLeuThrProGluGluAsnSerLysIle 160
 DB 657 CACAGAGACTTAAGCCCGGAAACCTGCTTTACCTTACCCCTGAGAGAACTCTAAGATC 716
 QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
 DB 717 ATGATCACTGACTTGTCTGTCTCAAGATGGAACAGAAATGGCATCATGCTCCCTGT 776
 QY 181 GlyThrProGlyTrpValAlaProGluValLeuAlaGlnLysProTrpSerLysAlaVal 200
 DB 777 GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCCAAGAAACCTTACAGCAAGGCTGTG 836
 QY 201 AspCysTrpSerIleGlyValIleThrTrpIleLeuLeuCysGlyTrpProPheTrp 220

DB 837 GATTGCTGGTCCATCGGCGTCATCACCTACATATTGCTGTGTGATACCCCATCTTAT 896
 QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTrpTrpGluPheGluSer 240
 DB 897 GAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAGTTTGAGTCT 956
 QY 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
 DB 957 CCATCTCGGATGACATTTCTGAGTCAGCCAGGACTTTATTGGCCACTTGCCTGAGAAG 1016
 QY 261 AspProAsnGluArgTrpThr-CysGluLysAlaLeuSerHisProTrpIleAspGlyAs 280
 DB 1017 GATCCGAACGAGGGGTAAACACTGTGAGAAGGGCTTGAGTCATCCCTGGATTGACGGTAA 1076
 QY 280 nThrAlaLeuHisArgAspIleTrpProSerVal-SerLeuGlnIleGlnLysAsnPheA 300
 DB 1077 CACAGCCCTCCACCGGACATCTACCCATCAGTCAGGCGTCCAGATCCAGAAACTTTG 1136
 QY 300 LaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValVal-HisHisMetArgLys 319
 DB 1137 CTAGAGCAAGTGGAGGCAAGCCTTCAACGACGAGCTGTGGTAGCACACAATGAGGAAG 1196
 QY 320 LeuHisMetAsnLeuHisSer-ProGlyValArgProGluValGluAsnArgProProG 339
 DB 1197 CTACACATGAACCTGCACAGCCCGGGCGTCCGCCAGAGGTGGAGAACAGGCGCGCTGA 1256
 QY 339 uThrGlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaPr 359
 DB 1257 AACTCAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGTAGATCACCATCACGAGGCACC 1316
 QY 359 oValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLeu---ProCysGlnHisG 378
 DB 1317 TGTCTCGACACACAGTGTAGCACTCCCTGCTGTGACCCCAATATACCTGCCAGCATCG 1376
 QY 378 yArgArgProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuH 398
 DB 1377 CCGCGGCGCCACTGCCCTTGGTGGCAGGTCCCTCACTGCTGCTCAATGGCTCCCTCCA 1436
 QY 398 sIleSerSerSerLeuValProMetHisGlnGlySerLeuAlaGlyProCysGlyCy 418
 DB 1437 CATCAGCAGCAGCCTGGTGGCCATGCATAGGGGTCCCTGGCGCGGCGCTGTGGCTG 1496
 QY 418 sCysSer-SerCysLeuAsnIleGlySerLysGlyLysSerSerTrpCysSerGluPro 438
 DB 1497 CTGCTCCAAGCTGCTGAACATTTGGAGCAAAAGAAAGTCTCTCTACTGTCTGAGCCCA 1556
 QY 438 hrLeuLeuLysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProVal 458
 DB 1557 CACTCTCAAAAGAGCCCAACAAAACAGAACTTCAAGTCGAGGTCTATGTTACCGGTTA 1616
 QY 458 yAlaSerGly-SerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
 DB 1617 AAGCCAGTGGCATGCTTCCACTGCGGCGAGGCGAGACTGGAGTCTGTCTCATTTATG 1673

RESULT 5

AAD18817
 ID AAD18817 standard; cDNA; 1736 BP.

XX AAD18817;

DT 18-DEC-2001 (first entry)

XX Human kinase (PKIN)-2 cDNA.

DE Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
 XX cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
 KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
 KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
 KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;
 KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
 KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
 KW asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
 KW antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.

XX	OS	Homo sapiens.		DB:	22	Gaps:	11
XX	PH	Key	Location/Qualifiers	US-09-960-643-2 (1-476) x AAD18817 (1-1736)			
XX	CD5	159..1232					
FT	FT	/tag= a		Qy	1	MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg	20
FT	FT	/product= "Human PKIN-2 protein"		Db	159	ATGCCCGGGAGAACGGCAGAGCAGCTCTCTCTGGAAAAAGCAAGCTGAAGACATCAAG	218
FT	FT	159..278					
FT	FT	/tag= b		Qy	21	LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal	40
FT	FT	279..1229		Db	219	AAGATCTTCGAGTTCAAAGAGACCTCGAAGCCGGGCCCTTTCCGAAGTGGTTTAGCT	278
FT	FT	/tag= c					
FT	FT	/product= "Mature human PKIN-2 protein"		Qy	41	LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla	59
XX	PN	WO200181555-A2.		Db	279	GAAGAGAAGCACTGGCAAGCTCTTGTCTGTGAAGTGTATCCTTAAAGAGCGCTGAAG	338
XX	PD	01-NOV-2001.					
XX	XX	20-APR-2001; 2001WO-US12992.		Qy	60	PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysLysLysLysGluAsn	79
XX	XX	20-APR-2000; 2000US-199021P.		Db	339	GGCAAGGAAAGCAGCATAGAGATGAGATAGCCGCTCTGAGAAAGATTAAAGCATGAAAT	398
PR	PR	28-APR-2000; 2000US-200226P.		Qy	80	IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu	99
PR	PR	05-MAY-2000; 2000US-202339P.		Db	399	ATTGTTGCCCTGGGAAGACATTTATGAAAGCCCAATCACCTGACTTGGTCATGCACTG	458
PR	PR	11-MAY-2000; 2000US-203505P.					
PR	PR	18-MAY-2000; 2000US-205564P.		Qy	100	ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp	119
PR	PR	26-MAY-2000; 2000US-207739P.		Db	459	GTGTCGGGTGAGAGCTGTTTGACCGGATAGTGAGAGAGGGGTTTATACAGAGAGGAT	518
XX	XX	01-JUN-2000; 2000US-208795P.					
XX	PA	(INCY-) INCYTE GENOMICS INC.		Qy	120	AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle	139
PI	PI	Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;		Db	519	GCCAGCATCTGTATGCCCAAGTCTTGAGCGCGCTCTACTATCTCCACAGAATGGGCATC	578
PI	PI	Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Walia NK, Yao MG;		Qy	140	ValHisArgAspLeuLysProGluAsnLeuLeuTyrThrProGluGluAsnSerLys	159
PI	PI	Patterson C, Burrill JB, Marcus GA, Zingler KA, Recipon SA, Lu Y;		Db	579	GTCCACAGACCTCAAGCCCGAAATCTCTTGACTACAGTCAAGTCAAGAGGAGTCCAA	638
PI	PI	Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;		Qy	160	IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr	178
PI	PI	Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;		Db	639	ATAATGATCATGACTGTTGGATTGTTCAAAAATGGAGGGCAAGAGATGTGATCCACT	698
PI	PI	Gururajan R;		Qy	179	AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys	198
DR	DR	WPI: 2001-611740/70.		Db	699	GCCTGTGGAATCCAGGCTATGCTGCTCTGAAATCCTCGCCAGAAACCTTACAGCAA	758
DR	DR	P-PSDB; AAEL1768.		Qy	199	AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro	218
XX	XX	Human kinases and nucleic acids, useful for preventing diagnosing and		Db	759	GCCGTTGACTGCTGGTCCATCGGAGTGATGCTCTGCTCGCGCTACCTCCT	818
XX	XX	treating cancers, inflammation and immune disorders -		Qy	219	PheTyrGluGluThrGluSerLysLeuPheGluLysLysLysGluGlyTyrTyrGluPhe	238
XX	XX	Claim 5; Page 151-152; 166pp; English.		Db	819	TTTTATGATGAAATGACTCCAAGCTCTTTGAGCAGATCCTCAAGCGGGAATATGAGTTT	878
CC	CC	The present invention relates to human kinases (PKIN) and the nucleic		Qy	239	GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu	258
CC	CC	acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is		Db	879	GACTCTCCCTACTGGGATGACATCTCGACTCTGCAAAAGACTTCATTCGGAACCTGATG	938
CC	CC	used in the prevention, diagnosis and treatment of diseases cancers,		Qy	259	GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrPheAsp	278
CC	CC	adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,		Db	939	GAGAAGGACCCGATAAAAGATACACGCTGTGAGCAGCAGCTCGGCACCCATCGATCGCT	998
CC	CC	acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,		Qy	279	GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn	298
CC	CC	gout, microbial infections, cardiovascular disease and/or inflammation,		Db	999	GGTGACAGCCCTCAACAAAAACATCCAGAGTCCGTCAGCCCGCAGATCCGGAANAAC	1058
CC	CC	myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial		Qy	299	PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg	318
CC	CC	infarction, cataract, growth and development disorder, seizure disorder,		Db	1059	TTTGCCCAAGAGCAAAATGGAGACAGCATTTAATGCCACGCCCTGTGAGACATATGAGA	1118
CC	CC	pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage		Qy	319	LysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArg-----	336
CC	CC	disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.		Db	1119	AAACTACACCTCGCAGCAGCCCTGGACAGTTCAAATGCAAGTGTTCGACGACCTCAGT	1178
CC	CC	PKIN may be used to treat disorders associated with decreased PKIN		Qy	337	---ProProGluThrGlnAlaSerGluThrSerArgProSerProGluIleThrIle	355
CC	CC	expression by rectifying mutations or deletions in a patient's genome					
CC	CC	that affect the activity of PKIN by expressing inactive proteins or to					
CC	CC	supplement the patients own production of PKIN. PKIN nucleic acids may be					
CC	CC	used to produce the PKIN polypeptide, by inserting the nucleic acids into					
CC	CC	a host cell and culturing the cell to express the protein. PKIN nucleic					
CC	CC	acid and its complementary sequences may also be used as DNA probes in					
CC	CC	diagnostic assays to detect and quantitate the presence of similar					
CC	CC	nucleic acid sequences in samples and therefore which patients may be in					
CC	CC	need of restorative therapy. The present sequence is human PKIN-2 cDNA.					
XX	XX	Sequence 1736 BP; 463 A; 456 C; 466 G; 351 T; 0 other;					
SQ							
Alignment Scores:							
Pred. No.:	1.22e-97	Length:	1736				
Score:	1250.50	Matches:	272				
Percent Similarity:	62.85%	Conservative:	63				
Best Local Similarity:	51.03%	Mismatches:	110				
Query Match:	49.76%	Indels:	90				


```
Db 541 GCCTGTGGAATCCAGGCTATGCTCGCTCCTGAAGTCTCGCCAGAAACCTTACAGCAA 600
QY 199 AlavalaspysrtrpserlleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 601 GCCCTTACACTGCTGGTCCATCGGAGTGATGCTACATCTTGTCTCGCGCTACCTCCT 660
QY 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrGluPhe 238
Db 661 TTTTATGATGAAATGACTCCAAAGCTCTTGACGACATCTCAAGCGGGAATATGAGTTT 720
QY 239 GluSerProPheTrpAspPilleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 721 GACTCTCCCTACTGGGATGACATCTCGGACTCTGCAAAAGACTTCAATTCGGAACCTGATG 780
QY 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
Db 781 GAGAAGACCCGGAATAAAGATACAGTGTGACGAGCAGCTCGGCAACCATGGATCGCT 840
QY 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGluIleGlnLysAsn 298
Db 841 GGTGACACGCCCTCAACAAAACATCCAGAGTCCGTGAGCGCCCGGATCCGGAATAAC 900
QY 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
Db 901 TTTGCCAAGAGCAATGGAGACAAGCATTTAATGCCACCGCGCTGCTGAGACATATGAGA 960
QY 319 LysLeuHisMet 322
Db 961 AAACACTACACTC 972

RESULT 7
AAH25119
ID AAH25119 standard; DNA; 1074 BP.
XX
AC AAH25119;
XX
DT 22-AUG-2001 (first entry)
XX
DE Nucleotide sequence of a human kinase polypeptide.
XX
KW Human; kinase; human disease; human disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1074
FT /tag= a
FT /transl_except= "(pos: 838..843, aa: Ala)"
FT /product= "kinase"
XX
PN W0200142435-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-US33240.
XX
PR 07-DEC-1999; 99US-0169428.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Donoho G, Scoville J, Turner CA, Friedrich G, zambrowicz B;
PI Abuin A, Sands AT;
XX
DR WPI; 2001-381667/40.
DR P-PSDB; AAB84360.
XX
PT Novel isolated human kinase polynucleotide that shares structural
PT similarity with animal kinases including calcium/calmodulin-dependent
PT protein kinases and serine/threonine protein kinases, useful in
PT therapeutics -
XX
PS Disclosure; Page 30-31; 32pp; English.
XX
```

CC The present sequence encodes a kinase polypeptide. The kinase polynucleotides and polypeptides are useful in therapeutic, diagnostic and pharmacogenetic applications. They are useful for the detection of mutant kinases, or inappropriately expressed kinases for the diagnosis of a disease or disorder. They are useful for screening for drugs (or high throughput screening of combinatorial libraries) effective in the treatment of symptomatic or phenotypic manifestations of that disease or disorder. The polynucleotide sequence is useful as a source of CC probes and primers, which can be used to screen libraries, isolate CC clones, and prepare cloning and sequencing templates.

SQ Sequence 1074 BP; 313 A; 258 C; 276 G; 227 T; 0 other;

Alignment Scores:

Pred. No.: 1.5e-97 Length: 1074
Score: 1246.00 Matches: 237
Percent Similarity: 85.49% Conservative: 40
Best Local Similarity: 73.15% Mismatches: 45
Query Match: 49.58% Indels: 2
Db: 22 Gaps: 2

US-09-960-643-2 (1-476) x AAH25119 (1-1074)

```
QY 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 1 ATGGCCCGGAGAACGGCGAGAGCAGCTCTCTCTGGAAAAAGCAAGCTGAAGACATCAAG 60
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 61 AAGATCTTCGAGTTCAAAGAGACCCCTCGGAACCGGGGCCCTTTCGGAAGTGGTTTAGCT 120
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
Db 121 GAAGAGAAGGCAACTGGCAAGCTCTTCTGCTGAAGTGATATCCCTAAGAAAGCGCTGAAG 180
QY 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 191 GCGAAGGAACACGACATAGAGAATGAGATAGCCCTCTGAGAAAGATTAGCATGAAAT 240
QY 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
Db 241 ATTGTGCGCTGGAAGACATTTATGAAAGCCAAATCACCTGTACTTGGTTCATCGAGCTG 300
QY 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
Db 301 GTGTCGCGTGGAGAGCTGTTTGACCGGATAGTGAGAGGGGGTTTTATACAGAGAAGGAT 360
QY 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
Db 361 GCCAGCACTCTGATCCGCCAAGTCTTGAGCGCGCTGTACTATCTCCACAGAAATGGGCATC 420
QY 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159
Db 421 GTCCACAGAGACCTCAAGCCGGAATACTCTTGTACTACAGTCAAGATGAGGAGTCCAAA 480
QY 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
Db 481 ATAATGATCATGTGACTTTGGATTGTCAAAATGAGGGGCAAGAGATGTGATGTCCACT 540
QY 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
Db 541 GCCTGTGGAACCTCCAGGCTATGCTCGCTCTGAGTCTCGCCAGAAACCTTACAGCAA 600
QY 199 AlavalaspysrtrpserlleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 601 GCCCTTACACTGCTGGTCCATCGGAGTGATGCTACATCTTGTCTCGCGCTACCTCCT 660
QY 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrGluPhe 238
Db 661 TTTTATGATGAAATGACTCCAAAGCTCTTGGAGGAGATCTCAAGCGGGAATATGAGTTT 720
QY 239 GluSerProPheTrpAspPilleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 721 GACTCTCCCTACTGGGATGACATCTCGGACTCTGCAAAAGACTTCAATTCGGAACCTGATG 780
```

Db 721 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTCTATTCGGAACTGATG 780

Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278

Db 781 GAGAAGAGCCCGAATAAAGATACAGTGTGAGCAGGACGCTCGGCACCCATGGATCGCT 840

Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298

Db 841 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCGCTCAGCGCCAGATCCGGAATAAC 900

Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValAlaValHisHisMetArg 318

Db 901 TTGTCCCAAGAGCAATGGAGACAAGCATTTAATGCCACGCGCTGCTGAGACATATGAGA 960

Qy 319 LysLeuHisMet 322

Db 961 AACTACACCTC 972

RESULT 8

AAH25118

ID AAH25118 standard; DNA; 1158 BP.

XX

AC AAH25118;

XX

DT 22-AUG-2001 (first entry)

XX

DE Nucleotide sequence of a human kinase polypeptide.

XX

KW Human; kinase; human disease; human disorder; ss.

XX

OS Homo sapiens.

XX

FH Key

FT 1..1158

FT CDS

FT /*tag= a

FT /product= "kinase"

XX

PN W0200142435-A2.

XX

PD 14-JUN-2001.

XX

PF 07-DEC-2000; 2000WO-US33240.

XX

PR 07-DEC-1999; 99US-0169428.

XX

PA (LEXI-) LEXICON GENETICS INC.

XX

PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;

PI Abuln A, Sands AT;

XX

DR WPI; 2001-381667/40.

DR P-PSDB; AAB84359.

XX

PT Novel isolated human kinase polynucleotide that shares structural

PT similarity with animal kinases including calcium/calmodulin-dependent

PT protein kinases and serine/threonine protein kinases, useful in

PT therapeutics -

XX

PS Claim 1; Page 29; 32pp; English.

XX

CC The present sequence encodes a kinase polypeptide. The kinase

CC polynucleotides and polypeptides are useful in therapeutic, diagnostic

CC and pharmacogenic applications. They are useful for the detection of

CC mutant kinases, or inappropriately expressed kinases for the diagnosis

CC of a disease or disorder. They are useful for screening for drugs (or

CC high throughput screening of combinatorial libraries) effective in the

CC treatment of symptomatic or phenotypic manifestations of that disease

CC or disorder. The polynucleotide sequence is useful as a source of

CC probes and primers, which can be used to screen libraries, isolate

CC clones, and prepare cloning and sequencing templates.

XX

SQ Sequence 1158 BP; 323 A; 283 C; 304 G; 248 T; 0 other;

Alignment Scores:

Pred. No.: 1.67e-97 Length: 1158

Score: 1246.00 Matches: 237

Percent Similarity: 85.49% Conservative: 40

Best Local Similarity: 73.15% Mismatches: 45

Query Match: 49.58% Indels: 2

DB: 22 Gaps: 2

US-09-960-643-2 (1-476) x AAH25118 (1-1158)

Qy 1 MetGlyArgLysGluLysAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20

Db 1 ATGCCCGGGAGACGGCGAGAGCAGCTCTCTCCGAAAAAGCAAGCTGAAGCATCAAG 60

Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40

Db 61 AAGATCTTCGAGTTCAAAAGAGACCTCGGAACCGGGGCTTTTCCGAAGTGTGTTAGCT 120

Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59

Db 121 GAAGAGAAAGCAACTGGCAAGCTCTTCTGCTGAAGTGTATCCCTAAGAGGCGCTGAAG 180

Qy 60 PheArgAspSerSerLeuLysGlnIleAlaValLeuLysLysIleLysHisGluAsn 79

Db 181 GCGAAGAAACAGCATAGAGATGAGATAGCCGCTCTGAGAAAGATTAACCATGAAAT 240

Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99

Db 241 ATTGTTGCCCTGGAAAGACATTTATGAAGCCCAATCACCTGTACTTGTGGTCATGAGCTG 300

Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119

Db 301 GTGTCCGCTGGAGAGCTGTTGACCGGATAGTGGAAAGGGGTTTATACAGAGAAGAT 360

Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139

Db 361 GCCAGCACTCTGATCCGCCAAGTCTTGGACCGCTGTACTATCTCCACAGAATGGCATC 420

Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159

Db 421 GTCCACAGAGACTCAAGCCGAAAAATCTCTGTACTACAGTCAAGATGAGAGATCCAAA 480

Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178

Db 481 ATAATGATCAGTGACTTTGGATTGTCAAAATGGAGGCAAGAGAGATGTGATGCCACT 540

Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198

Db 541 GCCTGTGGAACCTCCAGGCTATGTCGCTCTCTGAAGTCTCTGCCAGAAACCTTACAGCAA 600

Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuLysGlyTyrProPro 218

Db 601 GCGGTGACTGCTGCTGCTCATCGAGTGTGCTACATCTTCTGCTGCGGTACCTCTCT 660

Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238

Db 661 TTTTATGATGAAATGACTTCAAGCTCTTTGAGCAGATCTCTCAAGCGCGATATGAGTTT 720

Qy 239 GluSerProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258

Db 721 GACTCTCCCTACTGGATGACATCTCCGACTCTGCAAAAGACTTCTCATTCGGAACCTGATG 780

Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278

Db 781 GAGAAGAGCCCGAATAAAGATACACGTTGAGCAGGAGCTCGGCGCCATGGATCGCT 840

Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298

Db 841 GTGTACACAGCCCTCAACAAAAACATCCACGAGTCGCTCAGCGCCAGATCCGGAATAAC 900

Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValAlaValHisHisMetArg 318

Db 901 TTGTCCCAAGAGCAATGGAGACAAGCATTTAATGCCACGCGCTGCTGAGACATATGAGA 960

QY 319 LysLeuHisMet 322
 Db 961 AACTACACCTC 972
 RESULT 9
 ID ABL60905
 XX ABL60905 standard; cDNA; 1565 BP.
 AC ABL60905;
 XX
 DT 23-SEP-2002 (first entry)
 DE Human Cam kinase I 39.05 polypeptide encoding cDNA.
 XX
 KW Human; Cam kinase I 39.05; nervous disease; arrhythmia; tumour;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 18..1085
 FT /*tag= a
 FT /product= "human Cam kinase I 39.05"
 XX
 CN1333357-A.
 XX
 PD 30-JAN-2002.
 XX
 PF 07-JUL-2000; 2000CN-0119405.
 XX
 PR 07-JUL-2000; 2000CN-0119405.
 XX
 PA (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI: 2002-305611/35.
 DR P-PSDB; ABB08178.
 XX
 PT Novel polypeptide human Cam kinase I 39.05 and polynucleotide for
 encoding said polypeptide -
 XX
 PS Claim 6; Page 26-27 (disclosure); 35pp; Chinese.
 XX
 CC The present invention discloses a novel polypeptide-human Cam kinase I
 CC 39.05, polynucleotide for coding the polypeptide and method for producing
 CC this polypeptide by DNA recombination technology. The polypeptide is
 CC useful for treating diseases such as nervous disease, arrhythmia, tumour
 CC and growth development disturbance disease. The present sequence
 CC represents the human Cam kinase I 39.05 polypeptide encoding cDNA.
 XX
 SQ Sequence 1565 BP; 408 A; 394 C; 410 G; 353 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 2.56e-97 Length: 1565
 Score: 1246.00 Matches: 237
 Percent Similarity: 85.49% Conservative: 40
 Best Local Similarity: 73.15% Mismatches: 45
 Query Match: 49.58% Indels: 2
 DB: 24 Gaps: 2
 US-09-960-643-2 (1-476) x ABL60905 (1-1565)
 QY 1 MetGlyArgLysGluGluAspCysSerSerTrpLysGlnThrThrAsnIleArg 20
 Db 18 ATGGCCGGGAGAACGGGAGCAGCTCCCTCGAAAGAACGAGCTGAAGACATCAAG 77
 QY 21 LysThrPheIlePheMetGluValLysGlySerGlyAlaPheSerGluValPheIleVal 40
 Db 78 AGATCTTCGAGTTCAAGAGACCTCGAACCGGGCTTTCCGAGTGGTTTACGT 137
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59

Db 138 GAAGAGAGGCAACTGGCAAGCTCTTTGCTGTGAAGTGTATCCCTAAGAGGCGCTGAAG 197
 QY 60 PheArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
 Db 198 GGCAGGAAAGCAGCATAGAGATGAGATAGCCCTCCTGAGAAGATTAAAGCATGAAAT 257
 QY 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
 Db 258 ATTGTTGCCCTGGAAAGACATTTATGAAAGCCCAATACACTGTACTTGGTCATGCAGCTG 317
 QY 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
 Db 318 GTGTCCGGTGGAGAGCTGTTTGGCCGATAGTGAGAGGGGTTTATACAGAGAAGGAT 377
 QY 120 AlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
 Db 378 GCCAGCACTCTGATCCGCCAAGTCTTGGAGCGCGTACTATCTCCACAGAATGGGCATC 437
 QY 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159
 Db 438 GTCCACAGAGACCTCAAGCCGAAATCTCTGTACTACAGTCAAGATGAGGATCCAAA 497
 QY 160 IleMetIleThrAspPheGlyLeuSerLysMetGluLysAsnGly---IleMetSerThr 178
 Db 498 ATATGATCATGTGACTTTGGATTGTCAAAATGGAGGGCAAGGAGATGTGTCCACT 557
 QY 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
 Db 558 GCCTGTGGAACCTCAGGCTATGTCGCTCCTGGAAGTCTGCCCCAGAAACCTTACACAAA 617
 QY 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
 Db 618 GCGTGTGACTGCTGGTCCATCGGAGTGATGCTTACATCTTGTCTGCGGCTACCCCTCT 677
 QY 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
 Db 678 TTTTATGATGAAAATGACTCCCAAGCTCTTTGACACATCTCAAGCGGGAATATGAGTTT 737
 QY 239 GluSerProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
 Db 738 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAGACTTTCATTCGGAACCTGATG 797
 QY 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
 Db 798 GAGAAGGACCCGAATAAAGATACACGTGTGAGCAGGCGACTCGGCACCCATGGATCGCT 857
 QY 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
 Db 858 GGTGACACAGCCCTCAACAAAAACATCCAGAGTCCGTGAGCCCGCCAGATCCGGAANAAC 917
 QY 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
 Db 918 TTTTCCCAAGAGCAATGGAGACAAGCATTTAATGCCACGCGCGCTGCTGAGACATATGAGA 977
 QY 319 LysLeuHisMet 322
 Db 978 AAATACACCTC 989
 RESULT 10
 ID AAD18826
 XX AAD18826 standard; cDNA; 1661 BP.
 AC AAD18826;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human kinase (PKIN)-11 cDNA.
 XX
 KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
 KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
 KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
 KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;

KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;
KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW obesity; restorative therapy; immunomodulatory; vaccine; cardiovascular;
KW antimicrobial; cytostatic; antiinflammatory; asthma; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 159..1226

FT /*tag= a

FT /product= "Human PKIN-11 protein"

XX WO200181555-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12992.

XX 20-APR-2000; 2000US-199021P.

PR 28-APR-2000; 2000US-200226P.

PR 05-MAY-2000; 2000US-202339P.

PR 11-MAY-2000; 2000US-203505P.

PR 18-MAY-2000; 2000US-205564P.

PR 26-MAY-2000; 2000US-207739P.

PR 01-JUN-2000; 2000US-208795P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
PI Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;
PI Patterson C, Burrill JD, Marcus GA, Zingler KA, Recipion SA, Lu Y;
PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
PI Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;
PI Gururajan R;

XX WPI; 2001-611740/70.

DR P-PSDB; AAE11777.

XX Human kinases and nucleic acids, useful for preventing diagnosing and
PT treating cancers, inflammation and immune disorders -

XX Claim 5; Page 159; 166pp; English.

CC The present invention relates to human kinases (PKIN) and the nucleic
CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
CC used in the prevention, diagnosis and treatment of diseases cancers,
CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
CC gout, microbial infections, cardiovascular disease and/or inflammation,
CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC infarction, cataract, growth and development disorder, seizure disorder,
CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC PKIN may be used to treat disorders associated with decreased PKIN
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of PKIN by expressing inactive proteins or to
CC supplement the patients own production of PKIN. PKIN nucleic acids may be
CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
CC a host cell and culturing the cell to express the protein. PKIN nucleic
CC acid and its complementary sequences may also be used as DNA probes in
CC diagnostic assays to detect and quantitate the presence of similar
CC nucleic acid sequences in samples and therefore which patients may be
CC in need of restorative therapy. The present sequence is human PKIN-11
CC cDNA.

XX Sequence 1661 BP; 408 A; 452 C; 448 G; 353 T; 0 other;

Alignment Scores:

Pred. No.: 2,79e-97 Length: 1661
Score: 1246.00 Matches: 237
Percent Similarity: 85.49% Conservative: 40
Best Local Similarity: 73.15% Mismatches: 45

Query Match:	49.58%	Indels:	2
DB:	22	Gaps:	2
US-09-960-643-2 (1-476) x AAD18926 (1-1661)			
Qy 1 MetGlyArgLysGluGluAspAspCysSerTrpLysGlnThrThrAsnIleArg 20			
Db 159 ATGGCCGGGAGAACGGCGAGAGCAGCTCTCTCTGAAAAGACGATGAAGACATCAAG 218			
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40			
Db 219 AAGATCTTCAGTTCAAAGAGAGCCCTCGGAACGGGGCCCTTTCCGAAGTGGTTTTAGCT 278			
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59			
Db 279 GAAGAGAAGGCAACTGGCAAGACTCTTTGCTGTGAAGTGTATCCCTAAGAAGGCGCTGAAG 338			
Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79			
Db 339 GCGAAGGAAAGCAGCATAGAGATGAGATGAGCGTCTCGAGAAAGATTAAAGCATGAAAT 398			
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99			
Db 399 ATTGTTGCCCTGGAAGACATTTATGAAGCCCAATCACCTGTACTTGTGTATGTCAGCTG 458			
Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119			
Db 459 GTCTCCGGTGGAGAGCTGTTTGACCGGATAGTGGAGAGGGGTTTTATACAGAAAGGAT 518			
Qy 120 AlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139			
Db 519 GCCAGCACTGTGATCGCCCAAGTCTTGGACGCGCTACTATCTCCACAGAATGGGCATC 578			
Qy 140 ValHisArgAspLeuLysProGluAsnLeuTyrLeuThrProGluGluAsnSerLys 159			
Db 579 GTCCACAGAGACCTCAAGCCCGAAATCTCTGTACTACAGTCAAGATGAGGAGTCCAAA 638			
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178			
Db 639 ATAATGATCAGTCACCTTTGGATTGTCAAAATGGAGGGCAAGAGATGTGATGTCCACT 698			
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198			
Db 699 GCCTGTGGAACCTCCAGGCTATGTCGCTCTGAAAGTCTCGCCAGAAACCTTACAGCAA 758			
Qy 199 AlavalAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218			
Db 759 GCCTGTGACTGCTGGTCCATCGGAGTGATGCTACATCTTGCTCTCGCGCTACCCCTCT 818			
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238			
Db 819 TTTTATGATGAAATGACTCCCAAGCTCTTTTGAGCAGATCTCTCAAGCGGGAATATGAGTTT 878			
Qy 239 GluSerProPheThrAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258			
Db 879 GACTCTCCCTACTTGGATGACATCTCCGACTCTGCAAAAGACTTCATTCGGAACCTGATG 938			
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278			
Db 939 GAGAAGGACCCGAATAAAAGATTACAGTGTGTGAGCAGGAGCTCGGCACCCATGGATCGCT 998			
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298			
Db 999 GGTGACACAGCCCTCAACAAAACATCCACGAGTCCGTGAGCGCCAGATCCGGAAAAAC 1058			
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318			
Db 1059 TTGTCGCAAGAGCAAAATGGAGACAGCATTTAATGCCCGCGCTCGTGAGACATATGAGA 1118			
Qy 319 LysLeuHisMet 322			
Db 1119 AAATACACCTC 1130			

RESULT 11
 AAH25120
 ID AAH25120 standard; DNA; 1671 BP.
 XX
 AC
 AC
 AAH25120;
 XX
 DT 22-AUG-2001 (first entry)
 XX
 DE Nucleotide sequence of a human kinase polypeptide.
 XX
 KW Human; kinase; human disease; human disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200142435-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 07-DEC-2000; 2000MO-US33240.
 XX
 PR 07-DEC-1999; 99US-0169428.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
 PI Abuin A, Sands AT;
 XX
 DR WPI; 2001-381667/40.
 XX
 PT Novel isolated human kinase polynucleotide that shares structural
 PT similarity with animal kinases including calcium/calmodulin-dependent
 PT protein kinases and serine/threonine protein kinases, useful in
 PT therapeutics .
 XX
 PS Disclosure; Page 32; 32pp; English.
 CC
 CC The present sequence encodes a kinase polypeptide. The kinase
 CC polynucleotides and polypeptides are useful in therapeutic, diagnostic
 CC and pharmacogenetic applications. They are useful for the detection of
 CC mutant kinases, or inappropriately expressed kinases for the diagnosis
 CC of a disease or disorder. They are useful for screening for drugs (or
 CC high throughput screening of combinatorial libraries) effective in the
 CC treatment of symptomatic or phenotypic manifestations of that disease
 CC or disorder. The polynucleotide sequence is useful as a source of
 CC probes and primers, which can be used to screen libraries, isolate
 CC clones, and prepare cloning and sequencing templates.
 XX
 SQ Sequence 1671 BP; 453 A; 437 C; 448 G; 333 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,81e-97 Length: 1671
 Score: 1246.00 Matches: 274
 Percent Similarity: 64.35% Conservative: 60
 Best Local Similarity: 52.79% Mismatches: 97
 Query Match: 49.58% Indels: 91
 DB: 22 Gaps: 11
 US-09-960-643-2 (1-476) x AAH25120 (1-1671)
 QY 1 MetGlyArgLysGluGluAspAspCysSerTrpLysLysGlnThrThrAsnIleArg 20
 Db 167 ATGGCCCGGGAGACGCGAGACGCTCTCTCTGAAAGCAAGCTGAAGACATCAAG 226
 QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
 Db 227 AAGATCTTCGAGTCAAGAGACCTCTCGGAACCGGGCCCTTTCCGAAGTGGTTTAGCT 286
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
 Db 287 GAAGAGAAGCACTGCGCAAGCTTTTGTGCTGAAGTGTATCCCTTAAGAGGCGCTGAAG 346
 QY 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysGluAsn 79
 Db 1425 GAGGACTGTTTGGTTCAGGAGTGTAAAGACGAGCGCTGACCAACATGTGTAAACCCCATC 1484


```
|||||
Db 995 GAGAGGAGCCCAATAAAGATACAGTGTGAGCAGGAGCTGCGCACCCATGGATCGCT 1054
|||||
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
|||||
Db 1055 GGTGACACAGCCCTCAACAAAACATCCACGAGTCCGTGAGCGCCAGATCCGGAAGAAC 1114
|||||
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
|||||
Db 1115 TTGGCAGAGCAATGGAGACAGCATTTAATGCCAGCGCGCTGCTGAGACATATGAGA 1174
|||||
Qy 319 LysLeuHisMet 322
|||||
Db 1175 AAACATACACCTC 1186

RESULT 14
ID AAC90432 standard; cDNA; 1694 BP.
XX
AC AAC90432;
XX
DT 19-MAR-2001 (first entry)
XX
DE Murine Dendritic Cell Kinase 3 coding sequence.
XX
KW Murine; Dendritic Cell Kinase; MDCK-3; autoimmune disorder; allergy;
KW wound healing; periodontal disease; inflammatory disease; tumour;
KW infection; Ca2+/calmodulin-dependent kinase family; ss.
XX
OS Mus musculus.
XX
PN WO200073468-A1.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14696.
XX
PR 28-MAY-1999; 99US-0136781.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Bird TA, Virca GD, Martin U, Anderson DM;
XX
DR WPI; 2001-061546/07.
XX
PT P-PSDB; AAB50055.
XX
PT Novel murine and human kinase nucleic acids useful for treating
XX inflammations, infections, tumors, allergies, autoimmune diseases, and
XX for stimulating or suppressing immune responses -
XX
PS Claim 1; Page 86; 106pp; English.
XX
CC The present sequence is the coding sequence for Murine Dendritic Cell
CC Kinase 3 (MDCK-3). The protein encoded by the present sequence is useful
CC for treating a variety of disorders listed in the disclosure of the
CC specification, including autoimmune disorders, allergic reactions,
CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair
CC and replacement, burns, incisions and ulcers, periodontal disease,
CC inflammatory diseases, tumours and bacterial, viral or fungal infection.
CC MDCK-3 is a member of the Ca2+/calmodulin-dependent kinase family.
XX
SQ Sequence 1694 BP; 437 A; 444 C; 464 G; 349 T; 0 other;

Alignment Scores:
Pred. No.: 1,26e-96 Length: 1694
Score: 1238.50 Matches: 258
Percent Similarity: 68.74% Conservative: 52
Best Local Similarity: 57.21% Mismatches: 90
Query Match: 49.28% Indels: 51
DB: 22 Gaps: 9

US-09-960-643-2 (1-476) x AAC90432 (1-1694)
```

```
Qy 1 MetGlyArgLysGluGlnAspAspCysSerTrpLysLysGlnThrThrAsnIleArg 20
|||
Db 243 ATGCCCGGAGAGAGCGGAGAGCAGCTCTCTCGAAGAAAGCAAGACAGACATTAA 302
|||
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
|||
Db 303 AAGATCTTCAGTTCAGAGAGACCTCGGAACCTGGGCCCTTTCTGAAGTGTGTTAGCC 362
|||
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle--LysLysSerProAla 59
|||
Db 363 GAGAGAGAGCTACTGGGAAGCTCTTCAGTGAAGTGCATCCCGAAGAGGCGCTGAAG 422
|||
Qy 60 PheArgAspSerSerLeuGluAsnGlnIleAlaValLeuLysLysIleLysHisGluAsn 79
|||
Db 423 GGCAAGGAGAGCAGCATCGAAGAGAGATTGCCGTGCTTAGAAGATTAAAGCATGAAAC 482
|||
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
|||
Db 483 ATTGTTGCTTGGGAAGATATTTATGAAGCCCAATCACCTCTACCTGGTCTACGAACTT 542
|||
Qy 100 ValSerGlyGlyLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
|||
Db 543 GTGTCTGTGGAGAACTTTCGATCGGATAGTGGAGAGGGGTTTACACAGAGAAGAT 602
|||
Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
|||
Db 603 GCCAGCACTCTCATCCGCCAGGTCTCGATCGGTATCTATCTCCACAGAAATGGGCATT 662
|||
Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159
|||
Db 663 GTCCACAGGAGACCTCAAGCCGGAGATCTCTTATCTATCTACAGTCAAGACGAGGAGTCCAAA 722
|||
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
|||
Db 723 ATAATGATCAGTGACTTTGGCTTGTGCAAAATGGAGGCGCAAGAGATGTGATGCCAG 782
|||
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
|||
Db 783 GCCTGCGGAGACCCAGGCTATGTTGCTCCGGAAGTTCTGCCCAAGAACGTCACAGCAA 842
|||
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
|||
Db 843 GCTGTGACACTGCTGCTCCATCGGGTGATCGCTATATCTTGTCTGTGTGTACCTCTCT 902
|||
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
|||
Db 903 TTTTATGATGAAATGACTCGAAGCTCTTGAACAGATCTCTCAAGGCGCAAGATATGAGTTT 962
|||
Qy 239 GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
|||
Db 963 GATTCCCCCTACTGGGATGACATCTCCGACTCTGCCAAGACTTCATTCGGAATCTGATG 1022
|||
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrIleAsp 278
|||
Db 1023 GAGAAAGACCCAAATAAAGATACACTTGTGAGCGAGCGACCTCGACACCCATGGATTGCT 1082
|||
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
|||
Db 1083 GGTGACACAGCCCTTAGCAAAAACATTCAGAAATCTGTAGTCCCAAGATCCCGAAGAAAT 1142
|||
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArg 318
|||
Db 1143 TTTGCAAAAGAGCAAAATGGAGCAAGCGTTTAAACGCCAGCGGCTGTGAGACATATCGG 1202
|||
Qy 319 LysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProPro 338
|||
Db 1203 AGGCTCCAGCTT-----GGCAGCAGCTGGACAGCTTCAATGCAAGTGTCT 1247
|||
Qy 339 GluThrGlnAlaSerGluThrArgProSerSerProGluIleThrIle----- 355
|||
Db 1248 TCAAGCAACCTCAGTTTGGCCAGCAAAAGATTGTGCGTCTGCCACCTCCACGCTCTG 1307
|||
Qy 356 -----ThrGluAla 358
```

```

Db 1308 TAGTTTCCTTCTTCTTCGTCGGGGTTCGAGGATTCGGAGCTCGAGAGGAGACCCAGGCC 1367
      |||:::|
Qy 359 ProValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLeu-----Pro 374
      |||:::|
Db 1368 CACCACGTGTACAAACAGGCGCACACTGGAAGCAAGTACCGCGCTCGGAGGTGGAACCCA 1427
      |||:::|
Qy 375 CysGlnHisGlyArgProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsn 394
      |||:::|
Db 1428 GGGGGCAGGCGCGGGAAGAGAGCCCT---GGCCGAGGAGCTCCGCAATCAGAAC 1484
      |||:::|
Qy 395 GlySerLeuHisLeSerSerLeuValProMetHisGlnGlySerLeuAlaGly 414
      |||
Db 1485 -----CCACCCAC----- 1493
      |||
Qy 415 ProCysGlyCysCysSerSerCysLeuAsnIle 425
      |||
Db 1494 CTTGCATGGTGC-----ACCTGCATAGGACTG 1520
      |||
RESULT 15
AAD04775
ID AAD04775 standard; cDNA; 1578 BP.
XX
AC AAD04775;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human death domain-containing receptor (DDCR) cDNA from HDPV291 clone.
XX
KW Human; death domain-containing receptor; DDCR; immunosuppressive;
KW antiarthritic; antirheumatic; antiproliferative; cytostatic;
KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmological; gene therapy;
KW immunodeficiency disease; Acquired immune deficiency syndrome;
KW AIDS; leukemia; autoimmune disease; systemic lupus erythematosus;
KW hyperproliferative disorder; neoplasm; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; cardiovascular disorder;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW ocular disorder; corneal infection; degenerative disease; SMA; apoptosis;
KW spinal muscular atrophy; epithelial cell proliferation; infection;
KW cancer; wound healing; skin aging; chemotaxis; HDPV291 clone; ss.
XX
OS Homo sapiens.
XX
FH Key
FT 288..1130
FT CDS
FT /*tag= a
FT /product= "Human death domain-containing receptor"
XX
XX
XX WO200129063-A2.
XX
XX 26-APR-2001.
XX
XX 17-OCT-2000; 2000WO-US28666..
XX
XX 18-OCT-1999; 99US-0159585.
XX
XX 24-NOV-1999; 99US-0167246.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Ruben SM;
XX
XX WPI; 2001-300316/31.
XX
XX P-PSDB; AAE01034.
XX
XX New death-domain containing receptor polynucleotides and polypeptides,
XX useful for treating and diagnosing cancer -
XX
XX Claim 1; Page 283-284; 298pp; English.
XX
XX The patent discloses novel death domain-containing receptor (DDCR)
XX cDNAs and their corresponding proteins. DDCR cDNA and protein are
XX used to prevent, treat or ameliorate a medical condition in mammals.

```

They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. The DDCR protein and its antibodies are used in the diagnosis and treatment of disorders such as immunodeficiency diseases (e.g. Acquired immune deficiency syndrome (AIDS), leukemia) autoimmune diseases (e.g. systemic lupus erythematosus, rheumatoid arthritis), hyperproliferative disorders (e.g. neoplasms of the breast or liver), cerebrovascular disorders (e.g. cerebral ischaemia, angiogenesis), cardiovascular disorders (e.g. cardiac arrest), neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's disease), ocular disorders (e.g. corneal infection), degenerative diseases (e.g. spinal muscular atrophy-SMA), cancer, aberrant apoptosis, disorders of the placenta or uterus and infections caused by bacteria, viruses and fungi. The DDCR proteins are used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, to support cell culture of primary tissues, to regenerate tissues and in chemotaxis. The DDCR sequences are used in correcting aberrant cellular apoptosis by gene therapy.

The present sequence is human death domain-containing receptor CC (DDCR) cDNA from clone HDPV291 (ATCC No: PTA839).

XX Sequence 1578 BP; 432 A; 381 C; 401 G; 364 T; 0 other;

Alignment Scores:

Pred. No.:	1.54e-96	Length:	1578
Score:	1237.00	Matches:	235
Percent Similarity:	86.16%	Conservative:	39
Best Local Similarity:	73.90%	Mismatches:	42
Query Match:	49.22%	Indels:	2
DB:	22	Gaps:	2

US-09-960-643-2 (1-476) x AAD04775 (1-1578)

Qy	7	AspAspCysSerSerTrpLysLysGlnThrThrAsnIleArgLysThrPheIlePheMet	26
Db	15	GAGAGCAGCTCTCTCTGAAAAAGCAAGCTGAAGACATCAAGAGATCTTCGAGTTCAAA	74
Qy	27	GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly	46
Db	75	GAGACCTCGAACCGGGGCTTTTCGAAAGTGTTCGTAAGAGGCAAGCACTGC	134
Qy	47	LysLeuPheAlaLeuLysCysIle--LysLysSerProAlaPheArgAspSerSerLeu	65
Db	135	AAGCTCTTGTGTGAAGTGTATCCCTAAGAAGCGCTGAAGGCAAGGAAGACAGCAT	194
Qy	66	GluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIleValThrLeuAsp	85
Db	195	GAGAAATGAGATAGCCGCTCTGAGAAAGATTAAAGCATGAAATATTTGTCCTGGAAG	254
Qy	86	IleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuValSerGlyGlyGluLeu	105
Db	255	ATTATGAAGGCCCAAAATACCTGTACTTGTGTATGTCAGCTGTGCTGGAGAGCTG	314
Qy	106	PheAspArgIleLeuGluA:gglyValTyrThrGluLysAspAlaSerLeuValIleGln	125
Db	315	TTTGACCGGATAGTGGAGAGGGGTTTATACAGAGAAGGATGCCAGCACTCTGATCGC	374
Qy	126	GlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLys	145
Db	375	CAAGTCTTGGACGCGGTGTACTTCTCCAGAAATGGGCATCGTCACAGAGACCTCAAG	434
Qy	146	ProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIleMetIleThrAspPhe	165
Db	435	CCCGAAATCTCTGTACTACAGTCAAGATGAGAGAGTCCAAATATATGATCACTGACTT	494
Qy	166	GlyLeuSerLysMetGluGlnAsnGly---IleMetSerThrAlaCysGlyThrProGly	184
Db	495	GGATTGTCAAAATGAGAGGCAAGAGATGTGATGTCCACTGCCTGTGGAACTCCAGGC	554
Qy	185	TyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSer	204
Db	555	TATGCTGCTCTCTGAAGTCTCGGCCAGAAACCTTACAGCAAGGCCGTGTGACTGCTGTC	614

```
QY 205 ileGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyrGluGluThrGlu 224
Db 615 ATCGGAGTGATTGCTACATCTTGCTGCGCTACCTCTTTTATGATGAAATGAC 674
QY 225 SerLysLeuPheGluLysIleGlyTyrTyrGluPheGluSerProPheTrpAsp 244
Db 675 TCCAAGCTTTTGACGATCCTCAAGCGGGAATAGATTGACTCTCCCTACTGGAT 734
QY 245 AspileSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLysAspProAsnGlu 264
Db 735 GACATCTCGGACTCTGCAAAAGACTTCATTCCGAACCTGATGGAGAAGACCCGAATAAA 794
QY 265 ArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsnThrAlaLeuHis 284
Db 795 AGATACACGTGTGACGAGCAGCTCGGCACCCATGGATCGCTGGTGACACAGCCCTCAAC 854
QY 285 ArgAspileTyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrp 304
Db 855 AAAAAACATCCACGAGTCCGTCAGCGCCAGATCCGGAAAAACTTTGCCAAGAGCAATGG 914
QY 305 ArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeuHisMet 322
Db 915 AGACAGCATTTAATGCCACGCCCTCGTGACACATATGAGAAAACACTACACCTC 968
```

Search completed: March 15, 2003, 05:11:17
Job time : 344 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 17:30:16 ; Search time 182 Seconds
(without alignments)
9436.615 Million cell updates/sec

Title: US-09-960-643-1
Perfect score: 2447
Sequence: 1 tggagtggagctcaagcag.....ttttctctataaaaaaaa 2447

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 501302 seqs, 350932545 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1636	66.9	1738	9	US-09-935-464-4
2	1239	50.6	1383	9	US-09-935-464-2
3	903	36.9	157875	9	US-09-935-464-1
4	462	18.9	480	9	US-09-935-464-6
5	450	18.4	501	9	US-09-935-464-7
6	319	13.0	319	9	US-09-935-464-50
7	274	11.2	476	9	US-09-764-868-189
8	212	8.7	467	9	US-09-935-464-47
9	209	8.5	470	9	US-09-935-464-48
10	185	7.6	356	9	US-09-935-464-49
11	149	6.1	1361	10	US-09-731-872-177
12	80	3.3	386	9	US-09-935-464-6
13	32	1.3	474	9	US-09-796-692-7777
14	32	1.3	1372	10	US-09-817-181-1
15	27	1.1	27	9	US-09-935-464-45
16	25	1.0	25	10	US-09-971-118-5
17	25	1.0	31	10	US-09-971-118-6
18	23	0.9	23	9	US-09-935-464-43
19	23	0.9	23	9	US-09-935-464-44

20	23	0.9	296	9	US-10-114-893-103	Sequence 103, App
21	23	0.9	534	10	US-09-839-185-15	Sequence 15, Appl
c 22	22	0.9	22	9	US-09-935-464-56	Sequence 56, Appl
23	21	0.9	21	9	US-09-935-464-39	Sequence 39, Appl
24	21	0.9	21	9	US-09-935-464-53	Sequence 53, Appl
c 25	21	0.9	21	9	US-09-935-464-54	Sequence 54, Appl
c 26	20	0.8	20	9	US-09-935-464-17	Sequence 17, Appl
27	20	0.8	20	9	US-09-935-464-18	Sequence 18, Appl
c 28	20	0.8	358	10	US-09-770-791-570	Sequence 570, App
c 29	20	0.8	1108	10	US-09-770-445-96	Sequence 96, Appl
c 30	20	0.8	1148	10	US-09-778-844-206	Sequence 206, App
c 31	20	0.8	2550	10	US-09-985-675-8	Sequence 8, Appl
c 32	20	0.8	3307	9	US-09-158-722-5	Sequence 5, Appl
c 33	20	0.8	3785	10	US-09-223-490-9	Sequence 9, Appl
c 34	20	0.8	3919	10	US-09-985-675-7	Sequence 7, Appl
c 35	20	0.8	42450	10	US-09-815-048-3	Sequence 3, Appl
c 36	19	0.8	337	9	US-09-954-531-560	Sequence 560, App
c 37	19	0.8	452	10	US-09-770-444-621	Sequence 621, App
c 38	19	0.8	649	10	US-09-770-149-538	Sequence 538, App
39	19	0.8	1126	9	US-10-161-510-9	Sequence 9, Appl
40	19	0.8	1355	9	US-10-161-510-8	Sequence 8, Appl
41	19	0.8	1757	10	US-09-822-849A-313	Sequence 313, App
42	19	0.8	1892	12	US-10-078-929-105	Sequence 105, App
43	19	0.8	2085	10	US-09-925-301-222	Sequence 222, App
44	19	0.8	2255	9	US-09-954-531-624	Sequence 624, App
45	19	0.8	2258	10	US-09-964-824A-261	Sequence 261, App

ALIGNMENTS

RESULT 1
US-09-935-464-4
; Sequence 4, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Harrington-Martin, Rory
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY
; FILE REFERENCE: 3322/1H702 US1
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 1738
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-4

Query Match	66.9%;	Score 1636;	DB 9;	Length 1738;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1736;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy 53	CTTCAACTCTGGAGCAATGGGTGCGAAGGAAGAGATGACTGCTGCTGGAAGAAAC	112		
Db 1	CTTCAACTCTGGAGCAATGGGTGCGAAGGAAGAGATGACTGCTGCTGGAAGAAAC	60		
Qy 113	AGACCACACATCCGGAACCTTCATTTTATGGAAGTCTGGATCAGAGCTTTCT	172		
Db 61	AGACCACACATCCGGAACCTTCATTTTATGGAAGTCTGGATCAGAGCTTTCT	120		
Qy 173	CAGAAGTTTCTGGTGAAGCAAGACTGACTGGGAAGCTCTTCTGCTGAAGTGCATCA	232		
Db 121	CAGAAGTTTCTGGTGAAGCAAGACTGACTGGGAAGCTCTTCTGCTGAAGTGCATCA	180		
Qy 233	AGAAGTCACTGCTCCGGGACAGCAGCTGGAGTAAGTGTCTGTGTGAAAGA	292		
Db 181	AGAAGTCACTGCTCCGGGACAGCAGCTGGAGTAAGTGTCTGTGTGAAAGA	240		

QY 293 TCAGCATGAAACATTGTGACCCCTGGAGGACATCTATGAGAGACACCACCCTACTACC 352
Db 241 TCAAGCATGAAACATTGTGACCCCTGGAGGACATCTATGAGAGACACCACCCTACTACC 300
QY 353 TGGTCATGCAGCTGTGTTCTGTTGGGAGCTCTTTGACCGGATCTCGAGCGGGTGCT 412
Db 301 TGGTCATGCAGCTGTGTTCTGTTGGGAGCTCTTTGACCGGATCTCGAGCGGGTGCT 360
QY 413 ACACAGAGAAGATGCGAGTCTGGTGTATCCAGCAGTCTTTGCGAGTGAAATACCTAC 472
Db 361 ACACAGAGAAGATGCGAGTCTGGTGTATCCAGCAGTCTTTGCGCAGTGAAATACCTAC 420
QY 473 ATGAGATGCGATGCTCCAGAGACTTAAAGCCGAAACCTGCTTACCTTACCCCTG 532
Db 421 ATGAGATGCGATGCTCCAGAGACTTAAAGCCGAAACCTGCTTACCTTACCCCTG 480
QY 533 AAGAGAACTCTAAGATCATGATCACTGACTTTGGTCTGTCCAGATGGAACAGATGCA 592
Db 481 AAGAGAACTCTAAGATCATGATCACTGACTTTGGTCTGTCCAGATGGAACAGATGCA 540
QY 593 TCATGTCCACTGCTGCGGACCCAGGCTACGTGGCTCCAGAGTGTGGCCCAAGAAC 652
Db 541 TCATGTCCACTGCTGCGGACCCAGGCTACGTGGCTCCAGAGTGTGGCCCAAGAAC 600
QY 653 CCTACAGCAAGGCTGTGGATTGCTGCTTCCATCGGCTCATCACTACATATTGCTCTGTG 712
Db 601 CCTACAGCAAGGCTGTGGATTGCTGCTTCCATCGGCTCATCACTACATATTGCTCTGTG 660
QY 713 GATACCCCGGCTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAGATCAAGGAGGCT 772
Db 661 GATACCCCGGCTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAGATCAAGGAGGCT 720
QY 773 ACTATGATTGTAGTCTCCATTCTGGATGACATTTCTGAGTCAGCAAGGACTTTATT 832
Db 721 ACTATGATTGTAGTCTCCATTCTGGATGACATTTCTGAGTCAGCAAGGACTTTATT 780
QY 833 GCCACTTGTCTGAGAGGATCCGAAGCAGCGGTACACCTGTGAGAGGCTTTGAGTCATC 892
Db 781 GCCACTTGTCTGAGAGGATCCGAAGCAGCGGTACACCTGTGAGAGGCTTTGAGTCATC 840
QY 893 CCTGGATTGCGGAAACACAGCCCTCCACCGGAGACATCTACCACAGTCAGCTCCAGA 952
Db 841 CCTGGATTGCGGAAACACAGCCCTCCACCGGAGACATCTACCACAGTCAGCTCCAGA 900
QY 953 TCCAGAGAACTTTGCTTAAGAGCAAGTGGAGCAAGCTTCAACGCGAGCTGTGGTGC 1012
Db 901 TCCAGAGAACTTTGCTTAAGAGCAAGTGGAGCAAGCTTCAACGCGAGCTGTGGTGC 960
QY 1013 ACCACATGAGGAAGCTTACATGAACCTGACAGAGCCGCGGTCCGCCAGAGGTGGAGA 1072
Db 961 ACCACATGAGGAAGCTTACATGAACCTGACAGAGCCGCGGTCCGCCAGAGGTGGAGA 1020
QY 1073 ACAGCCGCTTGAACCTCAAGCTCAGAAACCTTAGCCAGCTCCCTGAGATCAGCA 1132
Db 1021 ACAGCCGCTTGAACCTCAAGCTCAGAAACCTTAGCCAGCTCCCTGAGATCAGCA 1080
QY 1133 TCACGAGGCACTGTCTCTGGACACAGTGTAGCACTTCCCTGCGCTGACCCCAATACCT 1192
Db 1081 TCACGAGGCACTGTCTCTGGACACAGTGTAGCACTTCCCTGCGCTGACCCCAATACCT 1140
QY 1193 GCCAGATGGCGCGCGCCACCTGCGCTGGTGGAGGTCCTTCAACTGCTGTTCAATG 1252
Db 1141 GCCAGATGGCGCGCGCCACCTGCGCTGGTGGAGGTCCTTCAACTGCTGTTCAATG 1200
QY 1253 GCTCCCTCCACATCAGCAGAGCTGTGTGCTTGCATGATCAGGSGTCCCTGGCCCGCGGC 1312
Db 1201 GCTCCCTCCACATCAGCAGAGCTGTGTGCTTGCATGATCAGGSGTCCCTGGCCCGCGGC 1260
QY 1313 CTTGTGCTGTCTGCTCAGCTGCTGNACATTGGGAGCAAGAAAGTCTCTACTGCT 1372
Db 1261 CTTGTGCTGTCTGCTCAGCTGCTGNACATTGGGAGCAAGAAAGTCTCTACTGCT 1320

QY 1373 CTGAGCCACACTCTCTCAAAAGGCCAACAAAAACAGAACTTCAAGTCGAGGTCTATGG 1432
Db 1321 CTGAGCCACACTCTCTCAAAAGGCCAACAAAAACAGAACTTCAAGTCGAGGTCTATGG 1380
QY 1433 TACCAGTTAAAGCCAGTGGCAGCTCCCACTCCCGGAGGAGCAGACTGGAGTCTGTCTCA 1492
Db 1381 TACCAGTTAAAGCCAGTGGCAGCTCCCACTCCCGGAGGAGCAGACTGGAGTCTGTCTCA 1440
QY 1493 TTATGTGATTCCTGGAGCTGTGCTATGCTCACTGCAATTTTCAGGAGACATATTCACT 1552
Db 1441 TTATGTGATTCCTGGAGCTGTGCTATGCTCACTGCAATTTTCAGGAGACATATTCACT 1500
QY 1553 CTTCTGCTCTTCCAAACCTGCTTATCCCGCAGAGAGGAGGAGGAGGAGGAGGAGG 1612
Db 1501 CTTCTGCTCTTCCAAACCTGCTTATCCCGCAGAGGAGGAGGAGGAGGAGGAGGAGG 1560
QY 1613 CAGGGCTTAGCAGGAGCTTTCTGGCCAGAGAACACAGCTGCTGCCAGGGGGAGCC 1672
Db 1561 CAGGGCTTAGCAGGAGCTTTCTGGCCAGAGAACACAGCTGCTGCCAGGGGGAGCC 1620
QY 1673 CCTCATAGGAGCCCGCAGGAGGAGCCCAAGGCTTAGAAGCTTTGTTGAAGCTGTGAGCA 1732
Db 1621 CCTCATAGGAGCCCGCAGGAGGAGCCCAAGGCTTAGAAGCTTTGTTGAAGCTGTGAGCA 1680
QY 1733 GGAGAGCGGTCCCAACAGCTTCAGGTCTCCCTGACCTGCCCTGCTATGCCCCAC 1790
Db 1681 GGAGAGCGGTCCCAACAGCTTCAGGTCTCCCTGACCTGCCCTGCTATGCCCCAC 1738

RESULT 2

US-09-935-464-2
; Sequence 2, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH.
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-2

Query Match 50.6%; Score 1239; DB 9; Length 1383;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1339; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 70 ATGGGTGGAAGGAAGAGATGACTGCTGGAAGAACACACACCAACATCCGG 129
Db 1 ATGGGTGGAAGGAAGAGATGACTGCTGGAAGAACACACACCAACATCCGG 60
QY 130 AAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTTCTCAGAAGTTTCTCTGGTG 189
Db 61 AAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTTCTCAGAAGTTTCTCTGGTG 120
QY 190 AAGCAAGACTGACTGGGAGCTCTTGTCTGAGTGCATCAAGAGTCACTGCGCTTC 249
Db 121 AAGCAAGACTGACTGGGAGCTCTTGTCTGAGTGCATCAAGAGTCACTGCGCTTC 180
QY 250 CGGACAGCAGCTGAGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCATGAAAAATT 309
Db 181 CGGACAGCAGCTGAGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCATGAAAAATT 240
QY 310 GTGACCTGGAGACATCTATGAGAGCACCCCACTACTACCTGGTCTATGCACTGTGT 369

Db 241 GTGACCTGGAGACATCTATGAGACACACCCACTACTACTGGTCAATGACGTTGTT 300
QY 370 TCTGGTGGGAGCTCTTTGACCGGATCCCTGGAGCGGGTGTCTACACAGAGAAGATGCC 429
Db 301 TCTGGTGGGAGCTCTTTGACCGGATCCCTGGAGCGGGTGTCTACACAGAGAAGATGCC 360
QY 430 AGTCTGGTGTATCCAGCAGGTCTTGTGCGGAGTGAATACCTACATGAGAATGGCATCGTC 489
Db 361 AGTCTGGTGTATCCAGCAGGTCTTGTGCGGAGTGAATACCTACATGAGAATGGCATCGTC 420
QY 490 CACAGAGACTTAAGCCCGAAACCTGCTTACCTTACCCCTGAAGAACTCTAAGATC 549
Db 421 CACAGAGACTTAAGCCCGAAACCTGCTTACCTTACCCCTGAAGAACTCTAAGATC 480
QY 550 ATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATGTCCTACCTGCCGT 609
Db 481 ATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATGTCCTACCTGCCGT 540
QY 610 GGGACCCAGGCTACCTGGCTCCAGAACTGCTGGCCGCAAAACCCCTACAGCAAGCTGTG 569
Db 541 GGGACCCAGGCTACCTGGCTCCAGAACTGCTGGCCGCAAAACCCCTACAGCAAGCTGTG 600
QY 670 GATTGCTGGTCCATCGCGCTCATCACTACATATTGCTCTGTGGATACCCCTGCTAT 729
Db 601 GATTGCTGGTCCATCGCGCTCATCACTACATATTGCTCTGTGGATACCCCTGCTAT 560
QY 730 GAAGAAACGGAGTCTTAAGCTTTTCGAGAAAGATCAAGAGGGGTACTATGAGTTTGAAGTCT 789
Db 661 GAAGAAACGGAGTCTTAAGCTTTTCGAGAAAGATCAAGAGGGGTACTATGAGTTTGAAGTCT 720
QY 790 CCATTCTGGATCACATTTCTGAGTCAGCCAAAGGACTTTATTTGGCCACTTGTCTTGAGAG 849
Db 721 CCATTCTGGATCACATTTCTGAGTCAGCCAAAGGACTTTATTTGGCCACTTGTCTTGAGAG 780
QY 850 GATCCGAACAGCGGTACACCTGTGAGAAAGGCTTGTGAGTCATCCCTGGAATTGACGGAAC 909
Db 781 GATCCGAACAGCGGTACACCTGTGAGAAAGGCTTGTGAGTCATCCCTGGAATTGACGGAAC 840
QY 910 ACGGCCCTCCACGGGACATCTACCCATCAGTCAGGCTCCAGATCCAGAAAGACTTTGCT 969
Db 841 ACAAGCCCTCCACGGGACATCTACCCATCAGTCAGGCTCCAGATCCAGAAAGACTTTGCT 900
QY 970 AAGAGAAAGTGGAGGCAAGCTTCAACGACGAGCTGTGGTGCACCACATGAGGAAGTCA 1029
Db 901 AAGAGAAAGTGGAGGCAAGCTTCAACGACGAGCTGTGGTGCACCACATGAGGAAGTCA 960
QY 1030 CACATGAACCTGCACAGCCCGGCGTCCCGCCAGAGGTTGGAGACAGCGCGCTGAAACT 1089
Db 961 CACATGAACCTGCACAGCCCGGCGTCCCGCCAGAGGTTGGAGAAACAGGCGCGCTGAAACT 1020
QY 1090 CAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCACCGAGGACCTGTGTC 1149
Db 1021 CAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCACCGAGGACCTGTGTC 1080
QY 1150 CTGACCCACAGTGTAGCACTCCCTGCGCTGACCCATTAACCTGCGCAGCATGGCCGCGG 1209
Db 1081 CTGACCCACAGTGTAGCACTCCCTGCGCTGACCCATTAACCTGCGCAGCATGGCCGCGG 1140
QY 1210 CCCACTGCCCTGGTGGCAGGTCCCTCAACTGCTGGTCAATGGGTCCCTCCACATFACG 1269
Db 1141 CCCACTGCCCTGGTGGCAGGTCCCTCAACTGCTGGTCAATGGGTCCCTCCACATFACG 1200
QY 1270 AGCAGCTGGTGGCCATGATCAGGGGTCCCTGGCGCGCGGCTGTGGCTGCTGCTCC 1329
Db 1201 AGCAGCTGGTGGCCATGATCAGGGGTCCCTGGCGCGCGGCTGTGGCTGCTGCTCC 1260
QY 1330 AGCTGGCTGAACATTTGGGAGCAAGGAAAGTCTCTACTGCTGTGAGCCCACTCTCTC 1389
Db 1261 AGCTGGCTGAACATTTGGGAGCAAGGAAAGTCTCTACTGCTGTGAGCCCACTCTCTC 1320
QY 1390 AAAAGGCCCAACAAAAACAG 1410
|||||

Db 1321 AAAAAGGCCAACAAAAACAG 1341
RESULT 3
US-09-935-464-1
; Sequence 1, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 157875
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-1
Query Match 36.9%; Score 903; DB 9; Length 157875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 903; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1535 CAGGAGACATATTCAACTCCTCTCTCCAAACCTGGTGTCTATCCGGCAGAGGGAGG 1594
Db 147496 CAGGAGACATATTCAACTCCTCTCTCCAAACCTGGTGTCTATCCGGCAGAGGGAGG 147555
QY 1595 AAGCAGACGCAAGTGGAGCAGGGCTTAGCAGGAGCAGTTCCTGGCCAGAGCACCAGCCT 1654
Db 147556 AAGCAGACGCAAGTGGAGCAGGGCTTAGCAGGAGCAGTTCCTGGCCAGAGCACCAGCCT 147615
QY 1655 GCTGCCAGCGGGCAGCGCCCTCATAGGAGGCCAGGAGGAGGCCCAAGCGGTAGAAGCC 1714
Db 147616 GCTGCCAGCGGGCAGCGCCCTCATAGGAGGCCAGGAGGAGGCCCAAGCGGTAGAAGCC 147675
QY 1715 TTGTTGAAGCTGTGAGCAGAGAAAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGC 1774
Db 147676 TTGTTGAAGCTGTGAGCAGAGAAAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGC 147735
QY 1775 CTGCTCTATGCCCCACACCCCTACGTGCGGTCTGTGCGAGTGTAGTATAGTCTCTC 1834
Db 147736 CTGCTCTATGCCCCACACCCCTACGTGCGGTCTGTGCGAGTGTAGTATAGTCTCTC 147795
QY 1835 GCCTGGGTCTGTGTTGTCGTGAAAGCTTAATGGCTGGCCAGGCTGTGTCACCTT 1894
Db 147796 GCCTGGGTCTGTGTTGTCGTGAAAGCTTAATGGCTGGCCAGGCTGTGTCACCTT 147855
QY 1895 CTCCAAGCAAAAGCCATATGGAGCATCTACCAGACTCCCACTCTGCACACACTCACTCCC 1954
Db 147856 CTCCAAGCAAAAGCCATATGGAGCATCTACCAGACTCCCACTCTGCACACACTCACTCCC 147915
QY 1955 ACCTCTCAAGCCTTCAAGCTTGGCCAGATTTGGGCTCATTAATGTCGTCGCTGCCCAT 2014
Db 147916 ACCTCTCAAGCCTTCAAGCTTGGCCAGATTTGGGCTCATTAATGTCGTCGCTGCCCAT 147975
QY 2015 CTGCATGAATCAGAGGAGCTCCCATGGTGTGCTGTGAGCTCTTCAAGTTCTTAAT 2074
Db 147976 CTGCATGAATCAGAGGAGCTCCCATGGTGTGCTGTGAGCTCTTCAAGTTCTTAAT 148035
QY 2075 CCTTAACCTCCAGGATTAGCTCCCAAGTCGCTGAGACCCAGCCAGCAGCACTTCTTGGCCCT 2134
Db 148036 CCTTAACCTCCAGGATTAGCTCCCAAGTCGCTGAGACCCAGCCAGCAGCACTTCTTGGCCCT 148095
QY 2135 TCTCCCTGCCCTCAATCTAAAAGAGAGTGCCACACCCCTCCAAAGTGGATAGAAAGATTC 2194
Db 148096 TCTCCCTGCCCTCAATCTAAAAGAGAGTGCCACACCCCTCCAAAGTGGATAGAAAGATTC 148155

QY 2195 ATGAGTAAGGCTGCAAGGAATCTTATCTGCGCACATGTCTCCGTGCACACACCCAA 2254
|||||
Db 148156 ATGAGTAAGGCTGCAAGGAATCTTATCTGCGCACATGTCTCCGTGCACACACCCAA 148215
QY 2255 TGGAGTTAACCTTGGAAAGTTGACTATTTTAATGTCTGCCAGAGTTCTTAATCTGCGCCT 2314
|||||
Db 148216 TGGAGTTAACCTTGGAAAGTTGACTATTTTAATGTCTGCCAGAGTTCTTAATCTGCGCCT 148275
QY 2315 GTTCCCTTTTCTCTCTGAAAGTCCAGCACACCATTTCTGTCCCTTCCCGAGTTCCCTCG 2374
|||||
Db 148276 GTTCCCTTTTCTCTCTGAAAGTCCAGCACACCATTTCTGTCCCTTCCCGAGTTCCCTCG 148335
QY 2375 CCTCCACCCCTCCAGCTTCATGCTCAGTGTGTGCTTAATAAATGGACATATTTTCT 2434
|||||
Db 148336 CCTCCACCCCTCCAGCTTCATGCTCAGTGTGTGCTTAATAAATGGACATATTTTCT 148395
QY 2435 CTA 2437
|||
Db 148396 CTA 148398

RESULT 4
US-09-935-464-46
; Sequence 46, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US/09/935,464
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-46

Query Match 18.9%; Score 462; DB 9; Length 480;
Best Local Similarity 100.0%; Pred. No. 4.7e-218;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 CTCCTGAGGCAATGGTCCGAAAGGAAGATGACTGCAGTTCCTCGGAAGAACACAGACCA 118
|||||
Db 19 CTCCTGAGGCATGGTCCGAAAGGAAGATGACTGCAGTTCCTCGGAAGAACACAGACCA 78
QY 119 CCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAG 178
|||||
Db 79 CCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAG 138
QY 179 TTTTCTGCTGAGCAAGACACTGCTGGGAAGCTTCTGCTGAAGTGCATCAAGAAGT 238
|||||
Db 139 TTTTCTGCTGAGCAAGACACTGCTGGGAAGCTTCTGCTGAAGTGCATCAAGAAGT 198
QY 239 CACCTCCCTCCGGACACAGCCTGGGAATGAGATGAGATGCTGTGTGAAAAGATCAAGC 298
|||||
Db 199 CACCTCCCTCCGGACACAGCCTGGGAATGAGATGAGATGCTGTGTGAAAAGATCAAGC 258
QY 299 ATGAAAACATTTGACCCCTGGAGGACATCTATGAGACACACCCACTACTACCTGGTCA 358
|||||
Db 259 ATGAAAACATTTGACCCCTGGAGGACATCTATGAGACACACCCACTACTACCTGGTCA 318
QY 359 TGCAGCTTGTCTGCTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCTACACAG 418

Db 319 TGCAGCTTGTCTGCTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCTTACACAG 378
|||||
QY 419 AGAAGGATGCCAGTCTGATCCAGCAGCTCTGTCTCGCAGTGAATATACCTACATGAGA 478
|||||
Db 379 AGAAGGATGCCAGTCTGATCCAGCAGCTCTGTCTCGCAGTGAATATACCTACATGAGA 438
|||||
QY 479 ATGCATCGTCCACAGAGACTTAAAGCCCGAAACCTTCGCTTT 520
|||||
Db 439 ATGCATCGTCCACAGAGACTTAAAGCCCGAAACCTTCGCTTT 480

RESULT 5
US-09-935-464-7
; Sequence 7, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-7

Query Match 18.4%; Score 450; DB 9; Length 501;
Best Local Similarity 99.8%; Pred. No. 4e-212;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1525 CTGCAATTTTCAGGAGACATATTCAACTCCTCTCTCTTCCAAACCTGGTGTCTATCCGG 1584
|||||
Db 1 CTGCAATTTTCAGGAGACATATTCAACTCCTCTCTCTTCCAAACCTGGTGTCTATCCGG 60
QY 1585 CAGAGGGAGGAAGCAGAGCAAGTGGAGAGGGCTTAGCAGGAGCAGCTTTCTGGCCAGAA 1644
|||||
Db 61 CAGAGGGAGGAAGCAGAGCAAGTGGAGAGGGCTTAGCAGGAGCAGCTTTCTGGCCAGAA 120
QY 1645 GCACCAGCTCTGCCAGGGGCGAGCCCTCATAGAGGCCAGGAGGAGGCCCAAGG 1704
|||||
Db 121 GCACCAGCTCTGCCAGGGGCGAGCCCTCATAGAGGCCAGGAGGAGGCCCAAGG 180
QY 1705 CGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACCAGCTTCCAGGTCTC 1764
|||||
Db 181 CGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACCAGCTTCCAGGTCTC 240
QY 1765 CCTGACCTGCTGCTCTATGCCCCACACCTACGTGCGGTGGCTCTGTGAGTGTAGCTA 1824
|||||
Db 241 CCTGACCTGCTGCTCTATGCCCCACACCTACGTGCGGTGGCTCTGTGAGTGTAGCTA 300
QY 1825 GATAGCTCTGCGCTGGGTCTGTGCTTTGTGCTGAAAAGCTTAATGGGTGGCCAGGCT 1884
|||||
Db 301 GATAGCTCTGCGCTGGGTCTGTGCTTTGTGCTGAAAAGCTTAATGGGTGGCCAGGCT 360
QY 1885 GTGTCACCTTCTCCAAGCAAGCCATATGGAGCATCTACCCAGACTCCCCTACTTGCACAC 1944
|||||
Db 361 GTGTCACCTTCTCCAAGCAAGCCATATGGAGCATCTACCCAGACTCCCCTACTTGCACAC 420
QY 1945 ACTCAGTCCCACTCTCAAGCCTCAACCTCTTGCCAGATTTGGCTCATTATGTCGTT 2004
|||||
Db 421 ACTCAGTCCCACTCTCAAGCCTCAACCTCTTGCCAGATTTGGCTCATTATGTCGTT 480
QY 2005 GCCTGCCCATCTGCATGAATG 2025
|||||

```
Db 481 GCCTGCCATCTGCATGAATG 501

RESULT 6
US-09-935-464-50
; Sequence 50, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-50

Query Match 13.0%; Score 319; DB 9; Length 319;
Best Local Similarity 100.0%; Pred. No. 2.1e-147;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 CATCCTCAGAACTTCAACTCTGGAGCAATGGTGCAGGAAGAAAGATGACTGCAGTT 100
Db 1 CATCCTCAGAACTTCAACTCTGGAGCAATGGTGCAGGAAGAAAGATGACTGCAGTT 60
QY 101 CCTGGAGAAGAACAGACCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGAT 160
Db 61 CCTGGAGAAGAACAGACCAACATCCGGAACCTTCATTTTATGGAAGTCTGGGAT 120
QY 161 CAGAGCTTTCTCAGAAAGTTTCTGTGTGAAGCAAGACTGCTGGAGAGCTTTTGCTC 220
Db 121 CAGAGCTTTCTCAGAAAGTTTCTGTGTGAAGCAAGACTGCTGGAGAGCTTTTGCTC 180
QY 221 TGAAGTGCATCAAGAAGTCACTGCCCTCCGGACAGCAGCCCTGGAGATGAGATTGCTG 280
Db 181 TGAAGTGCATCAAGAAGTCACTGCCCTCCGGACAGCAGCCCTGGAGATGAGATTGCTG 240
QY 281 TGTGAAAAAGATCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCA 340
Db 241 TGTGAAAAAGATCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCA 300
QY 341 CCCACTACTACCTGGTCAAT 359
Db 301 CCCACTACTACCTGGTCAAT 319

RESULT 7
US-09-764-868-189
; Sequence 189, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PRZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 189
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-189

Query Match 11.2%; Score 274; DB 9; Length 476;
Best Local Similarity 100.0%; Pred. No. 3.7e-125;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGAGTGGAGCTCAAGCAGGATTTCCCGAGTCCCTGGGATCCTCAGAAGCTTCAACT 60
Db 145 TGGAGTGGAGCTCAAGCAGGATTTCCCGAGTCCCTGGGATCCTCAGAAGCTTCAACT 204
QY 61 CTGGAGGCAATGGGTGGAAGGAAGACTGACTGCAGTCTCTGGAAGAAACAGACCACC 120
Db 205 CTGGAGGCAATGGGTGGAAGGAAGACTGACTGCAGTCTCTGGAAGAAACAGACCACC 264
QY 121 AACATCCGGAACCTTCATTTTATGGAAGTGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180
Db 265 AACATCCGGAACCTTCATTTTATGGAAGTGTCTGGGATCAGGAGCTTTCTCAGAAGTT 324
QY 181 TTCCTGGTGAAGCAAGACTGACTGGGAAGCTTTTGTCTCTGAAGTGCATCAAGNAAGTCA 240
Db 325 TTCCTGGTGAAGCAAGACTGACTGGGAAGCTTTTGTCTCTGAAGTGCATCAAGNAAGTCA 384
QY 241 CCTGCTTCCGGACAGCAGCCTGGAGATGAGA 274
Db 385 CCTGCTTCCGGACAGCAGCCTGGAGATGAGA 418

RESULT 8
US-09-935-464-47
; Sequence 47, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 47
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-47

Query Match 8.7%; Score 212; DB 9; Length 467;
Best Local Similarity 99.4%; Pred. No. 1.6e-94;
Matches 312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 41 CATCCTCAGAAGCTTCAACTCTGGAGGCAATGGTGCAGGAAGAAAGATGACTGCAGTT 100
Db 1 CATCCTCAGAAGCTTCAACTCTGGAGGCAATGGTGCAGGAAGAAAGATGACTGCAGTT 60
QY 101 CTGGAAGAAACAGACCAACATCCGGAACCTTCATTTTATGGAAGTGTGGAT 160
Db 61 CTGGAAGAAACAGACCAACATCCGGAACCTTCATTTTATGGAAGTGTGGAT 120
QY 161 CAGGAGCTTTCTCAGAAGTTTCTGTGTGAAGCAAGACTGACTGGGAAGCTTTTGCTC 220
Db 121 CAGGAGCTTTCTCANAAGTTTCTGTGTGAAGCAAGACTGACTGGGAAGCTTTTGCTC 180
QY 221 TGAAGTGCATCAAGAAGTCACTTCCGGACAGCAGCCTGGAGATGAGATTGCTG 280
```

Db 181 TGAAGTCATCAAGAAAGTCACCTGCCTTCGGGACAGCAGCCTGNAGAAATGAGATTGCTG 240
Qy 281 TGTGAAAAAGATCAAGCATGAAACATTTGTGACCCCTGGAGGACATCTATGAGAGCACCA 340
Db 241 TGTGAAAAAGATCAAGCATGAAACATTTGTGACCCCTGGAGGACATCTATGAGAGCACCA 300
Qy 341 CCCACTACTACCTG 354
Db 301 CCCACTACTACCTG 314

RESULT 9
US-09-935-464-48
; Sequence 48, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE REFERENCE: 3322/14702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 48
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-48

Query Match 8.5%; Score 209; DB 9; Length 470;
Best Local Similarity 100.0%; Pred. No. 4.8e-93;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 183 CTGCTGGAAGCAAGACTGACTGGGAGCCTTTGCTCTGGAAGTCATCAAGAAAGTCAC 242
Db 146 CTGCTGGAAGCAAGACTGACTGGGAGCCTTTGCTCTGGAAGTCATCAAGAAAGTCAC 205
Qy 243 TGCCTTCCGGGACAGCAGCCTGGAGATGCTGTGTTGAAAAAGATCAAGCATGA 302
Db 206 TGCCTTCCGGGACAGCAGCCTGGAGATGCTGTGTTGAAAAAGATCAAGCATGA 265
Qy 303 AACATTGTGACCTGGAGCATCTATGAGAGCACCACCTACTACTACCTGGTCATGCA 362
Db 266 AACATTGTGACCTGGAGCATCTATGAGAGCACCACCTACTACTACTACCTGGTCATGCA 325
Qy 363 GCTGTTTCTGGTGGGAGCTCTTTGACC 391
Db 326 GCTGTTTCTGGTGGGAGCTCTTTGACC 354

RESULT 10
US-09-935-464-49
; Sequence 49, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE REFERENCE: 3322/14702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300

; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 49
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-49

Query Match 7.6%; Score 185; DB 9; Length 356;
Best Local Similarity 100.0%; Pred. No. 3.4e-81;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 183 CTGCTGGAAGCAAGACTGACTGGGAGCCTTTGCTCTGGAAGTCATCAAGAAAGTCAC 242
Db 146 CTGCTGGAAGCAAGACTGACTGGGAGCCTTTGCTCTGGAAGTCATCAAGAAAGTCAC 205
Qy 243 TGCCTTCCGGGACAGCAGCCTGGAGATGCTGTGTTGAAAAAGATCAAGCATGA 302
Db 206 TGCCTTCCGGGACAGCAGCCTGGAGATGCTGTGTTGAAAAAGATCAAGCATGA 265
Qy 303 AACATTGTGACCTGGAGCATCTATGAGAGCACCACCTACTACTACTACTGGTCATGCA 362
Db 266 AACATTGTGACCTGGAGCATCTATGAGAGCACCACCTACTACTACTACTGGTCATGCA 325
Qy 363 GCTTG 367
Db 326 GCTTG 330

RESULT 11
US-09-731-872-177
; Sequence 177, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 177
; LENGTH: 1361
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 65..1024
US-09-731-872-177

Query Match 6.1%; Score 149; DB 10; Length 1361;
Best Local Similarity 100.0%; Pred. No. 2.3e-63;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2299 TTCTAATCCCTGCCTCTGTTCCCTTTTCTCTCCCTTAAAGTCCAGACACACCATTTGTGTC 2358
Db 1177 TTCTAATCCCTGCCTCTGTTCCCTTTTCTCTCCCTTAAAGTCCAGACACACCATTTGTGTC 1236
Qy 2359 TTCCCCAGTTTCTCGCCCTCCACCCCTCAGCTTCATGCTCAGTGTGTGCTTAATAAA 2418
Db 1237 TTCCCCAGTTTCTCGCCCTCCACCCCTCAGCTTCATGCTCAGTGTGTGCTTAATAAA 1296
Qy 2419 ATGGACATATTTTCTCTAAAAA 2447

Db 1297 ATGCACATATTTCTCTCTAAAAAAAAA 1325
|||||

RESULT 12

US-09-935-464-6
; Sequence 6, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 386
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n = a or g or c or t/u
US-09-935-464-6

Query Match 3.3%; Score 80; DB 9; Length 386;
Best Local Similarity 100.0%; Pred. No. 2.6e-29; Mismatches 0; Indels 0; Gaps 0;
Matches 80; Conservative 0;

QY 1523 CACTGCAATTTTCAGGAGACATATTCAACTCTCTCTTCCAAACCTGGTCTATCC 1582
|||||

Db 3 CACTGCAATTTTCAGGAGACATATTCAACTCTCTCTTCCAAACCTGGTCTATCC 62

QY 1583 GGCAGAGGGAGGAGGCAGA 1602
|||||

Db 63 GGCAGAGGGAGGAGGCAGA 82

RESULT 13

US-09-796-692-7777/C
; Sequence 7777, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950

; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7777
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (19)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7777

Query Match 1.3%; Score 32; DB 9; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.4e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 32; Conservative 0;

QY 652 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 683
|||||

Db 229 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 198

RESULT 14

US-09-817-181-1
; Sequence 1, Application US/09817181
; Patent No. US20020142427A1
; GENERAL INFORMATION:
; APPLICANT: MERRILOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL001189
; CURRENT APPLICATION NUMBER: US/09/817,181
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Human
US-09-817-181-1

Query Match 1.3%; Score 32; DB 10; Length 1372;
Best Local Similarity 100.0%; Pred. No. 1.5e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 32; Conservative 0;

QY 652 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 683
|||||

Db 582 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 613

RESULT 15

US-09-935-464-45
; Sequence 45, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY
; FILE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA

```

; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 27
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA probe
US-09-935-464-45

```

```

Query Match      1.1%; Score 27; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 293 TCAAGCATGAAACATTGTGACCCCTGG 319
Db 1 TCAAGCATGAAACATTGTGACCCCTGG 27

```

```

Search completed: March 15, 2003, 05:09:04
Job time : 401 secs

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:03:36 ; Search time 38 Seconds
(without alignments)
1669.139 Million cell updates/sec

Title: US-09-960-643-2
Perfect score: 2513
Sequence: 1 MGRKEDECSNKKQTNTNR.....VKASGSHCRAGTGVCLIM 476

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2513	100.0	476	23	Human calmodulin k
2	2513	100.0	497	22	Human polypeptide
3	2499.5	99.5	567	22	Human polypeptide
4	1382	55.0	298	22	Human diagnostic a
5	1246	49.6	355	22	Human kinase (PKIN
6	1246	49.6	355	22	Human polypeptide
7	1246	49.6	355	23	Human Cam kinase I
8	1246	49.6	357	22	Human kinase (PKIN
9	1246	49.6	365	22	Human polypeptide
10	1246	49.6	365	22	Amino acid sequenc

11	1241	49.4	357	22	AAU03508	Human protein kina
12	1233	49.1	355	22	AAU03505	Murine Dendritic C
13	1231.5	49.0	356	22	AAU03505	Amino acid sequenc
14	1145	45.6	389	21	AAU03505	Amino acid sequenc
15	1070	42.6	358	19	AAU03505	Calmodulin-depende
16	1039	41.3	343	22	AAU03511	Human protein kina
17	1036	41.2	503	22	ABG05970	Novel human diago
18	990	39.4	343	22	ABG05970	Disease associated
19	957.5	38.1	405	22	ABG05970	Drosophila melanog
20	944.5	37.6	280	22	AAE01034	Human death domain
21	747.5	29.7	501	22	AAE01034	Human kinase (PKIN
22	747.5	29.7	501	22	AAE01034	Amino acid sequenc
23	747.5	29.7	501	22	AAE01034	Human protein kina
24	746.5	29.7	501	22	AAE01034	Human polypeptide
25	745.5	29.7	470	22	AAE01034	Human polypeptide,
26	745.5	29.7	470	22	AAE01034	Human polypeptide,
27	745.5	29.7	470	22	AAE01034	Human protein phos
28	745	29.6	470	22	AAE01034	Human polypeptide
29	733.5	29.2	501	22	AAE01034	Human polypeptide,
30	726.5	28.9	205	22	AAE01034	Human protein sequ
31	691.5	27.5	502	21	AAE01034	A rat calcium/calim
32	650.5	25.9	424	18	AAE01034	PSKH-1 serine kina
33	643.5	25.6	414	22	AAE01034	Human protein SEQ
34	643.5	25.6	482	22	AAE01034	Human protein SEQ
35	639.5	25.4	449	22	AAE01034	Novel human diago
36	625.5	24.9	766	23	AAE01034	Human PKIN-14 prot
37	625.5	24.9	766	23	AAE01034	Human serine/threo
38	625	24.9	264	20	AAE01034	Human protein kina
39	615.5	24.5	641	21	AAE01034	Amino acid sequenc
40	611.5	24.3	600	22	AAE01034	Human protein kina
41	611.5	24.3	600	22	AAE01034	Human kinase (PKIN
42	603	24.0	474	22	AAE01034	Novel human diago
43	599.5	23.9	493	22	AAE01034	Drosophila melanog
44	599.5	23.9	493	22	AAE01034	Drosophila melanog
45	599.5	23.9	493	22	AAE01034	Drosophila melanog

ALIGNMENTS

RESULT 1

AAE22764	AAE22764 standard; Protein; 476 AA.
ID	AAE22764 standard; Protein; 476 AA.
XX	
AC	AAE22764;
XX	
DT	09-AUG-2002 (first entry)
XX	
DE	Human calmodulin kinase, CAMK-XI.
XX	
KW	Human: cytosolic; antisense gene therapy; screening; protein kinase;
KW	cancer; liver; colon; tumour; inflammation; arthritis; synovium; CAMK-XI;
KW	enzyme; calmodulin kinase.
OS	Homo sapiens.
XX	
PN	WO200224947-A2.
XX	
PD	28-MAR-2002.
XX	
PF	20-SEP-2001; 2001WO-IB02237.
XX	
PR	20-SEP-2000; 2000US-233999P.
PR	02-OCT-2000; 2000US-237419P.
PR	02-OCT-2000; 2000US-237423P.
PR	04-OCT-2000; 2000US-238558P.
XX	
XX	10-MAY-2001; 2001US-290555P.
PA	(KINE-) KINETEK PHARM INC.
PA	(UYBR-) UNIV BRITISH COLUMBIA.
XX	
PI	Yoganathan T, Delaney AD;
XX	

DR WPI: 2002-394145/42.
DR N-PSDB; AAD36140.
XX
PT Diagnosing cancer, comprises determining the upregulation of expression
PT of a nucleic acid sequence encoding a protein kinase or upregulation of
PT expression of the protein kinase, in the cancer -
XX
XX
PS Claim 1; Page 64-65; 87pp; English.
XX
CC The invention relates to a method for screening biologically active agent
CC that modulates cancer associated protein kinase function. The invention
CC also relates to a method for diagnosing cancer comprising determining the
CC upregulation of expression of a nucleic acid sequence encoding a protein
CC kinase. The method is useful for diagnosing cancer. A protein kinase is
CC useful for screening biological agents that modulate cancer associated
CC protein kinase function. Downregulating the activity of protein kinase is
CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon
CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy
CC derived tumours and inflammatory samples such as arthritic synovium, for
CC amplified DNA in the cell or increased expression of corresponding mRNA
CC or protein and is also useful to detect differences in expression levels
CC such as molecular weight, amino acid and nucleotide sequences between the
CC two cells. The present sequence is human calmodulin kinase, CAMK-XI.
XX
XX Sequence 476 AA;
Query Match 100.0%; Score 2513; DB 23; Length 476;
Best Local Similarity 100.0%; Pred. No. 4.8e-220;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRKEEDCSWKKQTNNIRKTFIFMEVLGSGAFSEVFLVKQLTKLPAKCIKSPAF 60
Db 1 MGRKEEDCSWKKQTNNIRKTFIFMEVLGSGAFSEVFLVKQLTKLPAKCIKSPAF 60
QY 61 RDSSLENEIAVLKIKHENIVTLEDIYESTTHYLVLMQLVSGGELFDRILRGVYTEKDA 120
Db 61 RDSSLENEIAVLKIKHENIVTLEDIYESTTHYLVLMQLVSGGELFDRILRGVYTEKDA 120
QY 121 SLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 180
Db 121 SLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 180
QY 181 GTPGVAPVLAQPKYSKAVDCNSIGVITVILLCGYPPEEFTESKLFKIKEGYEFES 240
Db 181 GTPGVAPVLAQPKYSKAVDCNSIGVITVILLCGYPPEEFTESKLFKIKEGYEFES 240
QY 241 PFWDIDISEAKDFCHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLIQKNFA 300
Db 241 PFWDIDISEAKDFCHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLIQKNFA 300
QY 301 KSKWRQAFNAAVVHHMRKLMNLHSPGVPEVENRPPETQASSETSRPSPETITTEAPV 360
Db 301 KSKWRQAFNAAVVHHMRKLMNLHSPGVPEVENRPPETQASSETSRPSPETITTEAPV 360
QY 361 LDHVSVALPALTQLPCQGRRTAPGGRSLNCLVNGSLHISLSSLPVPMHQSLAAGPCGCCS 420
Db 361 LDHVSVALPALTQLPCQGRRTAPGGRSLNCLVNGSLHISLSSLPVPMHQSLAAGPCGCCS 420
QY 421 SCLNIGSKGSSYCSPTLLKANKKONFKSEYVMPVKASGSSHCRAGQTGYCLIM 476
Db 421 SCLNIGSKGSSYCSPTLLKANKKONFKSEYVMPVKASGSSHCRAGQTGYCLIM 476
RESULT 2
AAM41547
ID AAM41547 standard; Protein: 497 AA.
XX
AC AAM41547;
XX
XX 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6478.
XX

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR N-PSDB; AAI60703.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6478; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity; chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity; cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX Sequence 497 AA;
Query Match 100.0%; Score 2513; DB 22; Length 497;
Best Local Similarity 100.0%; Pred. No. 5.1e-220;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRKEEDCSWKKQTNNIRKTFIFMEVLGSGAFSEVFLVKQLTKLPAKCIKSPAF 60
Db 22 MGRKEEDCSWKKQTNNIRKTFIFMEVLGSGAFSEVFLVKQLTKLPAKCIKSPAF 81
QY 61 RDSSLENEIAVLKIKHENIVTLEDIYESTTHYLVLMQLVSGGELFDRILRGVYTEKDA 120
Db 82 RDSSLENEIAVLKIKHENIVTLEDIYESTTHYLVLMQLVSGGELFDRILRGVYTEKDA 141
QY 121 SLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 180
Db 142 SLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 201

QY 181 GTEGYVAPEVLAQKPYSKAVDCWSIGVITYILLGYPPEYEETESKLFKEIKEGYEFBS 240
|||||
Db 202 GTEGYVAPEVLAQKPYSKAVDCWSIGVITYILLGYPPEYEETESKLFKEIKEGYEFBS 261
|||||
QY 241 PFWDIDISEAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSIQIKNFA 300
|||||
Db 262 PFWDIDISEAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSIQIKNFA 321
|||||
QY 301 KSKWRQAFNAAVVHHMRKLNHLHSPGVRPEVENRPPETQASSTRSPSPETITTEAPV 360
|||||
Db 322 KSKWRQAFNAAVVHHMRKLNHLHSPGVRPEVENRPPETQASSTRSPSPETITTEAPV 381
|||||
QY 361 LDHSHVALPALTOQPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHQGSLAAGPCGCCS 420
|||||
Db 382 LDHSHVALPALTOQPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHQGSLAAGPCGCCS 441
|||||
QY 421 SCLNIGSKGKSSYCEPTLLKANKKQNFSEVMVPVKAGSSHCRAGOTGVCLIM 476
|||||
Db 442 SCLNIGSKGKSSYCEPTLLKANKKQNFSEVMVPVKAGSSHCRAGOTGVCLIM 497
|||||
RESULT 3
AAM39761
ID AAM39761 standard; Protein; 567 AA.
XX
AC AAM39761;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2906.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WC200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI58917.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 2906; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 567 AA:
Query Match 99.5%; Score 2499.5; DB 22; Length 567;
Best Local Similarity 98.6%; Pred No. 1.le-218;
Matches 476; Conservative 0; Mismatches 0; Indels 7; Gaps 1;
QY 1 MGRKEEDCSSWKQTNIIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKSPAF 60
|||||
Db 85 MGRKEEDCSSWKQTNIIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKSPAF 144
|||||
QY 61 RDSLENETAVLKKIKHENIVTLEDIYESTTHYLVLMOLVSGGELEFRLERGVYTEKDA 120
|||||
Db 145 RDSLENETAVLKKIKHENIVTLEDIYESTTHYLVLMOLVSGGELEFRLERGVYTEKDA 204
|||||
QY 121 SLVIQVLSAVKYLHENGIVHRDLK-----PENLLYLTPEENSKIMITDFGLSKMEQN 173
|||||
Db 205 SLVIQVLSAVKYLHENGIVHRDLKVSRRSEWPENLLYLTPEENSKIMITDFGLSKMEQN 264
|||||
QY 174 GIMSTACGTPGYVAPVLAQKPYSKAVDCWSIGVITYILLGYPPEYEETESKLFKEIKE 233
|||||
Db 265 GIMSTACGTPGYVAPVLAQKPYSKAVDCWSIGVITYILLGYPPEYEETESKLFKEIKE 324
|||||
QY 234 GYEFESPFWDIDISEAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSL 293
|||||
Db 325 GYEFESPFWDIDISEAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSL 384
|||||
QY 294 QIQKNEAKSKWRQAFNAAVVHHMRKLNHLHSPGVRPEVENRPPETQASSTRSPSPPEI 353
|||||
Db 385 QIQKNEAKSKWRQAFNAAVVHHMRKLNHLHSPGVRPEVENRPPETQASSTRSPSPPEI 444
|||||
QY 354 TITEAPVLDHSHVALPALTOQPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHQGSLAA 413
|||||
Db 445 TITEAPVLDHSHVALPALTOQPCQHGRRPTAPGGRSLNCLVNGSLHISSSLVPMHQGSLAA 504
|||||
QY 414 GPCGCCSSCLNIGSKGKSSYCEPTLLKANKKQNFSEVMVPVKAGSSHCRAGOTGV 473
|||||
Db 505 GPCGCCSSCLNIGSKGKSSYCEPTLLKANKKQNFSEVMVPVKAGSSHCRAGOTGV 564
|||||
QY 474 LIM 476
|||
Db 565 LIM 567
RESULT 4
AAU19443
ID AAU19443 standard; Protein; 298 AA.
XX
AC AAU19443;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human diagnostic and therapeutic polypeptide (DITHP) #29.
XX
KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;
KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;
KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;
KW respiratory disorder.
XX
OS Homo sapiens.
XX

PD 01-NOV-2001.
XX 20-APR-2001; 2001WO-US12992.
PF
XX 20-APR-2000; 2000US-199021P.
PR 28-APR-2000; 2000US-200226P.
PR 05-MAY-2000; 2000US-202339P.
PR 11-MAY-2000; 2000US-203505P.
PR 18-MAY-2000; 2000US-205564P.
PR 26-MAY-2000; 2000US-207739P.
PR 01-JUN-2000; 2000US-208795P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H. Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
PI Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Walia NK, Yao MG;
PI Patterson C, Burrill JD, Marcus GA, Zingler KA, Recipon SA, Lu Y;
PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
PI Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;
PI Gururajan R;
XX
XX WPI: 2001-611740/70.
DR N-PSDB; AAD18826.
DR
XX Human kinases and nucleic acids, useful for preventing diagnosing and
XX treating cancers, inflammation and immune disorders -
PT
XX Claim 1; Page 137-138; 166pp; English.
XX
XX The present invention relates to human kinases (PKIN) and the nucleic
XX acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
XX used in the prevention, diagnosis and treatment of diseases cancers.
CC
CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
CC gout, microbial infections, cardiovascular disease and/or inflammation,
CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC infarction, cataract, growth and development disorder, seizure disorder,
CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC PKIN may be used to treat disorders associated with decreased PKIN
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of PKIN by expressing inactive proteins or to
CC supplement the patients own production of PKIN. PKIN nucleic acids may be
CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
CC a host cell and culturing the cell to express the protein. PKIN nucleic
CC acid and its complementary sequences may also be used as DNA probes in
CC diagnostic assays to detect and quantitate the presence of similar
CC nucleic acid sequences in samples and therefore which patients may be
CC in need of restorative therapy. The present sequence is human PKIN-11
XX protein.
XX
XX Sequence 355 AA;
SQ

Query Match 49.6%; Score 1246; DB 22; Length 355;
Best Local Similarity 73.1%; Pred. No. 9.4e-105;
Matches 237; Conservative 40; Mismatches 45; Indels 2; Gaps 2;
2;

Qy 1 MGRKEEDCCSWKKQTNIKRTFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPA 59
Db 1 MARENGSSSSWKKQAEIDIKRFEKTLGTGAFSEVFLVLAEEKATGKLFVAKCIPKKALK 60
Qy 60 FRDSSLENEIAVLKKIKHENIVTLEDIYESTTHYVLQVLVSGGELFDRLTIRGVVYTEKD 119
Db 61 GKESSENEIAVLKIKHENIVALEDIYESP NHLYLQVLVSGGELFDRLVKGFTTEKD 120
Qy 120 ASLVQIVQLSAVKYLPHENGIVHRDLKPENLLYLTPPENSKIMITDFGLSKMEQNG-IMST 178
Db 121 ASTLIQVLDVAVYVLRMGIVHRDLKPENLLYVSQDEESKIMISDFGLSKMEKGDMVMT 180
Qy 179 ACGTPGYVAPVLAQKPYKAVDCWSIGVITYILLCGPPFPYETSKLFEKIKGYEYF 238
Db 181 ACGTPGYVAPVLAQKPYKAVDCWSIGVITYILLCGPPFPYENDSKLFEKILKAEYF 240

Qy 239 ESPFWDDISESAKDFICHLEKDPNERVYTCERKALSHPWIDGNALHRLDIYPSVLSQIOKN 298
Db 241 DSPYWDIDISDAKDFIRNLMEKDPNKRKYTCERQAARHPWIAGDTALNKNIHESVSQAIRKN 300
Qy 299 FAKSKWRQAFNAAAVVHHMRKLLHM 322
Db 301 FAKSKWRQAFNATAVVVHHMRKLLHL 324
RESULT 6
AAM41268
ID AAM41268 standard; Protein; 355 AA.
XX
XX AAM41268;
AC
XX 22-OCT-2001 (first entry)
DT
XX Human polypeptide SEQ ID NO 6199.
DE
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
OS
XX WO200153312-A1.
PN
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000WO-US34263.
PF
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB; AAT60424.
DR
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX Example 2; SEQ ID NO 6199; 10078pp; English.
PS
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX SQ Sequence 355 AA; Query Match 49.6%; Score 1246; DB 22; Length 355; Best Local Similarity 73.1%; Pred. No. 9.4e-105; Matches 237; Conservative 40; Mismatches 45; Indels 2; Gaps 2;

QY 1 MGRKEEDCCSWKKQTNNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPA 59
DB 1 MARENGESSWKKQAEIDIKKIFEFKETLGTGAFSEVFLAEAEKATGKLFVAKCIPKKALK 60

QY 60 FRDSSLENEIAVLKKIKHENIVTLLEDIYESTHYLYVMQLVSGGELFDRILRGVYTEKD 119
DB 61 GKESSENEIAVLKKIKHENIVALEDIYESPNNHLYVMQLVSGGELFDRIVERGFTYTKD 120

QY 120 ASLVIQQVLSAVKYLHENGIVHRDLKPNLLYLTPPEENSKIMITDFGLSKMEQNG-IMST 178
DB 121 ASTLIROQLDAVYVYLRHMGIVHRDLKPNLLYYSQDESKIMISDFGLSKMEKGDMVST 180

QY 179 ACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLGYPFPFETESKLPKEIKEGYEF 238
DB 181 ACGTPGYVAPEVLAQKPYSKAVDCWSIGVIAIYILLGYPFPFYDENDSKLPFEQILKAEYEF 240

QY 239 ESPFWDDISAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIOKN 298
DB 241 DSPYWDIDSDSAKDFIRNLMEKDPNKRYTCEQARHPWIAAGDTALNKNIHESVSAQIRKN 300

QY 299 FAKSKWRQAFNAAVVHMRKLHM 322
DB 301 FAKSKWRQAFNATAVVRHMRKLHL 324

RESULT 7
ABB08178 ID ABB08178 standard; Protein: 355 AA.
XX AC ABB08178;
XX DT 23-SEP-2002 (first entry)
XX DE Human Cam kinase I 39.05 polypeptide.
XX KW Human; Cam kinase I 39.05; nervous disease; arrhythmia; tumour; enzyme.
XX KW Homo sapiens.
XX OS CN1333357-A.
XX PN 30-JAN-2002.
XX PD 07-JUL-2000; 2000CN-0119405.
XX PF 07-JUL-2000; 2000CN-0119405.
XX PR (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.
XX PA Mao Y, Xie Y;
XX PI WPI; 2002-305611/35.
XX DR N-PSDB; ABL60905.
XX XX Novel polypeptide human Cam kinase I 39.05 and polynucleotide for encoding said polypeptide -
XX PS Claim 1; Page 27-28 (disclosure); 35pp; Chinese.
XX CC The present invention discloses a novel polypeptide-human Cam kinase I 39.05, polynucleotide for coding the polypeptide and method for producing this polypeptide by DNA recombination technology. The polypeptide is useful for treating diseases such as nervous disease, arrhythmia, tumour and growth development disturbance disease. The present sequence represents the human Cam kinase I 39.05 polypeptide.

XX SQ Sequence 355 AA; Query Match 49.6%; Score 1246; DB 23; Length 355; Best Local Similarity 73.1%; Pred. No. 9.4e-105; Matches 237; Conservative 40; Mismatches 45; Indels 2; Gaps 2;

QY 1 MGRKEEDCCSWKKQTNNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPA 59
DB 1 MARENGESSWKKQAEIDIKKIFEFKETLGTGAFSEVFLAEAEKATGKLFVAKCIPKKALK 60

QY 60 FRDSSLENEIAVLKKIKHENIVTLLEDIYESTHYLYVMQLVSGGELFDRILRGVYTEKD 119
DB 61 GKESSENEIAVLKKIKHENIVALEDIYESPNNHLYVMQLVSGGELFDRIVERGFTYTKD 120

QY 120 ASLVIQQVLSAVKYLHENGIVHRDLKPNLLYLTPPEENSKIMITDFGLSKMEQNG-IMST 178
DB 121 ASTLIROQLDAVYVYLRHMGIVHRDLKPNLLYYSQDESKIMISDFGLSKMEKGDMVST 180

QY 179 ACGTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLGYPFPFETESKLPKEIKEGYEF 238
DB 181 ACGTPGYVAPEVLAQKPYSKAVDCWSIGVIAIYILLGYPFPFYDENDSKLPFEQILKAEYEF 240

QY 239 ESPFWDDISAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIOKN 298
DB 241 DSPYWDIDSDSAKDFIRNLMEKDPNKRYTCEQARHPWIAAGDTALNKNIHESVSAQIRKN 300

QY 299 FAKSKWRQAFNAAVVHMRKLHM 322
DB 301 FAKSKWRQAFNATAVVRHMRKLHL 324

RESULT 8
AAE11768 ID AAE11768 standard; Protein: 357 AA.
XX AC AAE11768;
XX DT 18-DEC-2001 (first entry)
XX DE Human kinase (PKIN)-2 protein.
XX KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout; cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome; AIDS; Addison's disease; microbial infection; inflammation; osteoporosis; atherosclerosis; cardiovascular disease; myocardial infarction; anaemia; myasthenia gravis; cirrhosis; cataract; growth and development disorder; seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder; lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease; asthma; obesity; restorative therapy; cytostatic; immunomodulatory; antimicrobial; cardiovascular; antiinflammatory; vaccine.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Peptide 1..40 /label= Signal_peptide
FT Domain 24..270 /note= "Protein kinase domain"
FT Domain 23..279 /note= "Eukaryotic protein kinase domain"
FT Protein 41..357 /label= Human_mature_PKIN-2_protein
XX WO200181555-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US12992.
XX PR 20-APR-2000; 2000US-199021P.
XX PR 28-APR-2000; 2000US-200226P.
XX PR 05-MAY-2000; 2000US-202339P.

CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.

XX Sequence 357 AA;

Query Match 49.4%; Score 1241; DB 22; Length 357;
Best Local Similarity 72.8%; Pred. No. 2.7e-104;
Matches 236; Conservative 40; Mismatches 46; Indels 2; Gaps 2;

QY 1 MGRKEEDCSSWKKQTNIIRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCI-KKSPA 59
DB 1 MARENGESSSSWKKQAEIDIKKIFEFKFTLGTGAFSEVLAEEKATGKLFVAKCIPKKALK 60
QY 60 FRDSSLENEIAVLKKIKHENIVTLEDIYESTHYLYVMOLVSGGELFDRIERGYYTEKD 119
DB 61 GKSSSENEIAVLKKIKHENIVALEDIYSPNHLIYVMOLVSGGELFDRIERGYYTEKD 120
QY 120 ASLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMQNG-INST 178
DB 121 ASTLIRQVLDVAVYLHRMGIVHRDLKPNLLYYSQDEESKIMISDFGLSKMEGKGDVNST 180
QY 179 ACSTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKGYEYF 238
DB 181 ACSTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGYPPFYDENSKLFEQILKAEYEF 240
QY 239 ESPFWDDISAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKN 298
DB 241 DSPYWDIDISAKDFIRNLMKDPNRYTCEQAARHPWIAGDTALNKNIHESVSQAIRKN 300
QY 299 FAKSKWRQAFNAAVVHHMRKLHM 322
DB 301 FEKSKWRQAFNATAVVRHMRKLHL 324

RESULT 12

AAB50055

ID AAB50055 standard; Protein; 355 AA.

XX AAB50055;

XX 19-MAR-2001 (first entry)

XX Murine Dendritic Cell Kinase 3.

XX Murine; Dendritic Cell Kinase; MDC3-3; autoimmune disorder; allergy;
XX wound healing; periodontal disease; inflammatory disease; tumour;
XX infection; Ca2+/calmodulin-dependent kinase family.

XX Mus musculus.

XX WO200073468-A1.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14696.

XX 28-MAY-1999; 99US-0136781.

XX (IMMUNEX CORP.

XX Bird TA, Virca GD, Martin U, Anderson DM;

XX WPI; 2001-061546/07.
DR N-PSDB; AAC90432.
XX Novel murine and human kinase nucleic acids useful for treating
PT inflammations, infections, tumors, allergies, autoimmune diseases, and
PT for stimulating or suppressing immune responses -
XX Claim 10; Page 92-94; 106pp; English.
XX The present sequence is Murine Dendritic Cell Kinase 3 (MDC3-3). This
CC protein is useful for treating a variety of disorders listed in the
CC disclosure of the specification, including autoimmune disorders, allergic
CC reactions, myeloid or lymphoid cell deficiencies, wound healing and
CC tissue repair and replacement, burns, incisions and ulcers, periodontal
CC disease, inflammatory diseases, tumours and bacterial, viral or fungal
CC infection. MDC3-3 is a member of the Ca2+/calmodulin-dependent kinase
CC family.
XX Sequence 355 AA;

Query Match 49.1%; Score 1233; DB 22; Length 355;
Best Local Similarity 72.5%; Pred. No. 1.4e-103;
Matches 235; Conservative 40; Mismatches 47; Indels 2; Gaps 2;

QY 1 MGRKEEDCSSWKKQTNIIRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCI-KKSPA 59
DB 1 MARENGESSSSWKKQAEIDIKKIFEFKFTLGTGAFSEVLAEEKATGKLFVAKCIPKKALK 60
QY 60 FRDSSLENEIAVLKKIKHENIVTLEDIYESTHYLYVMOLVSGGELFDRIERGYYTEKD 119
DB 61 GKSSSENEIAVLKKIKHENIVALEDIYSPNHLIYVMOLVSGGELFDRIERGYYTEKD 120
QY 120 ASLVIOQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMQNG-INST 178
DB 121 ASTLIRQVLDVAVYLHRMGIVHRDLKPNLLYYSQDEESKIMISDFGLSKMEGKGDVNST 180
QY 179 ACSTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKGYEYF 238
DB 181 ACSTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGYPPFYDENSKLFEQILKAEYEF 240
QY 239 ESPFWDDISAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKN 298
DB 241 DSPYWDIDISAKDFIRNLMKDPNRYTCEQAARHPWIAGDTALNKNIHESVSQAIRKN 300
QY 299 FAKSKWRQAFNAAVVHHMRKLHM 322
DB 301 FAKSKWRQAFNATAVVRHMRRLQL 324

RESULT 13

AAB84360

ID AAB84360 standard; Protein; 356 AA.

XX AAB84360;

XX 22-AUG-2001 (first entry)

XX Amino acid sequence of a human kinase polypeptide.

XX Human; kinase; human disease; human disorder.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 280

XX /note= "Ala encoded by GCTGGT"

XX WO200142435-A2.

XX 14-JUN-2001.

XX 07-DEC-2000; 2000WO-US33240.

```
XX PR 07-DEC-1999; 99US-0169428.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;
XX PI Abulin A, Sands AT;
XX WPI; 2001-381667/40.
DR N-PSDB; AAH25119.
XX Novel isolated human kinase polynucleotide that shares structural
PT similarity with animal kinases including calcium/calmodulin-dependent
PT protein kinases and serine/threonine protein kinases, useful in
PT therapeutics -
XX Claim 4; Page 31; 32pp; English.
XX The present sequence represents a kinase polypeptide. The kinase
CC polynucleotides and polypeptides are useful in therapeutic, diagnostic
CC and pharmacogenic applications. They are useful for the detection of
CC mutant kinases, or inappropriately expressed kinases for the diagnosis
CC of a disease or disorder. They are useful for screening for drugs (or
CC high throughput screening of combinatorial libraries) effective in the
CC treatment of symptomatic or phenotypic manifestations of that disease
CC or disorder. The polynucleotide sequence is useful as a source of
CC probes and primers, which can be used to screen libraries, isolate
CC clones, and prepare cloning and sequencing templates.
XX Sequence 356 AA:
Query Match 49.0%; Score 1231.5; DB 22; Length 356;
Best Local Similarity 72.8%; Pred. No. 2e-103;
Matches 236; Conservative 40; Mismatches 45; Indels 3; Gaps 3;
QY 1 MGRKEEDCSSKKKOTIVRTIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI-KKSPA 59
DB 1 MARENGESSSSKKQAEIDIKIFEKELGTGAFSEVVAEEKATGKLFVAKCIPKRALK 60
QY 60 FRDSSLENEIAVLKIKHENIVTLEDIYESTHYLLVMQLVSGGELFDRILRGVYTEKD 119
DB 61 GRESSIENEIAVLRIKHENIVALEDIYESPNHLYVMQLVSGGELFDRIVEKGFYTEKD 120
QY 120 ASLVLTQOVLAVKYLHENGIVHRDLKPENILYLPTEENSKIMTDFGLSKMEQNG-TMST 178
DB 121 ASLTLRQVLDAVYVYLRHNGIVHRDLKPENILLYYSQDEESKIMISDFGLSKMEGRGDVMT 180
QY 179 ACGTGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGPPPYEETESKLRKIKRGYEF 238
DB 181 ACGTGYVAPEVLAQPKYSKAVDCWSIGVIAIYLLCGPPPYDENDSKLFEQILKAYEF 240
QY 239 ESPFWDDISAKDPFICHLLEKDPNERYTCESKALSHPWIDGNTALHRDIYPSVSLQIQKN 298
DB 241 DSPYWDIDSDSAKDPFIRNMEKDPNRYTCESQAAHPWI-ADTALNKNIHESVSAQIRKN 299
QY 299 FAKSKWRQAFNAVVHMRKLHM 322
DB 300 FAKSKWRQAFNATAVVRHMRKLHL 323
RESULT 14
RAY68793
ID AAY68793 standard; Protein; 389 AA.
XX AC AAY68793;
XX DT 16-MAY-2000 (first entry)
XX DE Amino acid sequence of a human phosphorylation effector PHSP-25.
XX KW Human; phosphorylation effector; PHSP; proliferative disorder;
XX immune disorder; neuronal disorder.
```

```
OS Homo sapiens.
XX Key Location/Qualifiers
FH Modified-site 31 /note= "potential phosphorylation site"
FT Modified-site 56 /note= "potential phosphorylation site"
FT Domain 73..311 /note= "protein kinase catalytic domain"
FT Modified-site 96 /note= "potential phosphorylation site"
FT Modified-site 134 /note= "potential phosphorylation site"
FT Modified-site 148 /note= "potential phosphorylation site"
FT Modified-site 149 /note= "potential phosphorylation site"
FT Domain 152..208 /note= "potential phosphorylation site"
FT Modified-site 165 /note= "PTK core domain"
FT Region 172..184 /note= "potential phosphorylation site"
FT Modified-site 186 /note= "STK core catalytic motif"
FT Modified-site 201 /note= "potential phosphorylation site"
FT Modified-site 257 /note= "potential phosphorylation site"
FT Modified-site 283 /note= "potential glycosylation site"
FT Modified-site 301 /note= "potential phosphorylation site"
FT Modified-site 343 /note= "potential phosphorylation site"
FT Modified-site 358 /note= "potential glycosylation site"
FT Modified-site 364 /note= "potential phosphorylation site"
FT Modified-site 375 /note= "potential glycosylation site"
XX WQ200006728-A2.
XX 10-FEB-2000.
XX 28-JUL-1999; 99WO-US17132.
XX 28-JUL-1998; 98US-0123494.
PR 14-SEP-1998; 98US-0152814.
PR 14-OCT-1998; 98US-0173482.
PR 03-NOV-1998; 98US-0106889.
PR 19-NOV-1998; 98US-0109093.
PR 22-DEC-1998; 98US-0113796.
PR 12-JAN-1999; 99US-0173482.
PR 12-JAN-1999; 99US-0229005.
XX (INCY-) INCYTE PHARM INC.
XX Hillman JL, Lal P, Tang YT, Corley NC, Guegler KJ, Baughn MR;
PI Patterson C, Bandman O, Au-Young J, Gorgone GA, Yue H, Azimzal Y;
PI Reddy R, Lu DAM, Shih LL;
XX WPI; 2000-183125/16.
DR N-PSDB; AA246162.
XX New human phosphorylation effectors useful for the diagnosis, treatment
PT and prevention of proliferative, immune and neuronal disorders -
XX Claim 1; Page 112-113; 142pp; English.
XX AAY68769-95 and AAY68797-99 represent human phosphorylation effectors
CC (PHSP), designated PHSP1-PHSP31 (the protein sequence for PHSP28 is not
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 17:23:21 ; Search time 3882 Seconds
(without alignments)
10208.747 Million cell updates/sec

Title: US-09-960-643-1

Perfect score: 2447

Sequence: 1 tggagtgaggactcaagcag.....ttttctctataaaaaaaaa 2447

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba.*
2: em_esthum.*
3: em_estlin.*
4: em_estmu.*
5: em_estrov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	787	32.2	1126	BM547443	BM547443 AGENCOURT
2	762	31.1	809	BI821474	BI821474 603038366
3	710	29.0	740	BI824483	BI824483 603038855
4	684	28.0	1048	BI084101	BI084101 602869466
5	665	27.2	1129	BM807335	BM807335 AGENCOURT
c 6	656	26.8	966	BI084897	BI084897 602869466

7	650	26.6	731	13	BI753035	BI753035 603025844
8	636	26.0	742	13	BI818261	BI818261 603032510
9	577	23.6	812	13	BI772626	BI772626 603060879
10	551	22.5	1068	14	BM921532	BM921532 AGENCOURT
c 11	479	19.6	479	14	BQ102407	BQ102407 121C07.X
12	473	19.3	824	12	BG715920	BG715920 602676667
13	450	18.4	501	9	AL134342	AL134342 DKF2P5470
c 14	426	17.4	484	10	AW016039	AW016039 UT-H-B1Op
c 15	410	16.8	410	9	AA838372	AA838372 of29a11.s
16	408	16.7	573	14	BQ086330	BQ086330 121C07.Y
17	284	11.6	554	13	BI834635	BI834635 603090418
18	275	11.2	656	13	BI667965	BI667965 603292877
19	241	9.8	315	9	AA351937	AA351937 EST59855
20	139	5.7	537	17	AQ265651	AQ265651 CITBT-E1-
21	109	4.5	510	17	AQ041427	AQ041427 CIT-HSP-2
22	89	3.6	89	12	BG170935	BG170935 602323912
23	80	3.3	386	14	R05661	R05661 yf87d09.r1
24	71	2.9	818	12	BG293660	BG293660 602390529
25	71	2.9	905	14	BQ934044	BQ934044 AGENCOURT
c 26	70	2.9	273	12	BF950790	BF950790 RC3-NN019
27	68	2.8	188	17	AQ078706	AQ078706 CIT-HSP-2
c 28	59	2.4	306	9	AI713924	AI713924 UI-R-AC1-
c 29	59	2.4	439	10	AW522224	AW522224 UI-R-B00-
c 30	59	2.4	451	10	AW254051	AW254051 UI-R-BJ0-
c 31	59	2.4	555	10	AW251224	AW251224 UI-R-BJ0-
32	57	2.3	445	9	AI215131	AI215131 qp41f02.x
33	56	2.3	974	12	BG173948	BG173948 602334021
34	50	2.0	288	12	BF524803	BF524803 UI-R-AC1-
35	36	1.5	375	10	BB871015	BB871015 BB871015
c 36	35	1.4	118	10	AW122501	AW122501 UI-M-BH2-
c 37	35	1.4	204	10	AW521699	AW521699 UI-R-B00-
c 38	35	1.4	375	12	BE944047	BE944047 UI-M-BH3-
c 39	35	1.4	377	10	AW492153	AW492153 UI-M-BH3-
40	35	1.4	425	12	BF599348	BF599348 263103.MA
c 41	35	1.4	433	9	AI850410	AI850410 UI-M-BG1-
c 42	35	1.4	503	12	BG381458	BG381458 UI-R-CTO-
c 43	32	1.3	262	10	AW955600	AW955600 EST367670
c 44	32	1.3	286	14	T93936	T93936 ye06e06.r1
c 45	32	1.3	314	9	AA365669	AA365669 EST76508

ALIGNMENTS

RESULT 1
BM547443 1126 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6507688 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5724450
DEFINITION 5', mRNA sequence.
ACCESSION BM547443
VERSION BM547443.1 GI:18781213
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1126)
AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L18A12713 row: 1 column: 19
High quality sequence stop: 623.
Location/Qualifiers
1. .1126

FEATURES
source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5724450"
/clone_lib="NIH_MGC_125"
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."
BASE COUNT      299 a      286 g      244 t      6 others
ORIGIN
Query Match      32.2%; Score 787; DB 13; Length 1126;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 837; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 GGAGTGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACTC 61
Db 1 GGAGTGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACTC 60
Qy 62 TGGAGGCAATGGGTGCGAAAGGAAGAGATGACTGCTGCTGCTGGAAGAACAGACCACCA 121
Db 61 TGGAGGCAATGGGTGCGAAAGGAAGAGATGACTGCTGCTGCTGGAAGAACAGACCACCA 120
Qy 122 ACATCCGGAACCTTCATTTTATGGAAGTGTCTGGGATCAGAGCTTCTCAGAGTTT 181
Db 121 ACATCCGGAACCTTCATTTTATGGAAGTGTCTGGGATCAGAGCTTCTCAGAGTTT 180
Qy 182 TCCTGGTGAACCAAGACTGACTGGGAGCTCTTCTGCTGAGTGCATCAAGACTCAC 241
Db 181 TCCTGGTGAACCAAGACTGACTGGGAGCTCTTCTGCTGAGTGCATCAAGAGTCAAC 240
Qy 242 CTGCTCTCCGGGACAGCAGCTGGAGAAATGAGATGCTGTGTGTGTAAGAAAGATCAAGCATG 301
Db 241 CTGCTCTCCGGGACAGCAGCTGGAGAAATGAGATGCTGTGTGTGTAAGAAAGATCAAGCATG 300
Qy 302 AAACATTTGACCTCGGAGACATCTATGAGAGCACCACCTACTACTCTGGTTCATCG 361
Db 301 AAACATTTGACCTCGGAGACATCTATGAGAGCACCACCTACTACTCTGGTTCATCG 360
Qy 362 AGCTTGTCTCTGTTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGA 421
Db 361 AGCTTGTCTCTGTTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGA 420
Qy 422 AGGATGCCAGTCTGGTGATCCAGAGGTCTTGTGCGGAGTGAATACCTACATGAGAATG 481
Db 421 AGGATGCCAGTCTGGTGATCCAGAGGTCTTGTGCGGAGTGAATACCTACATGAGAATG 480
Qy 482 GCATCTGCCAGAGACTTAAAGCCGAAACCTGCTTACCTTACCTTACCTGGAAGAGAACT 541
Db 481 GCATCTGCCAGAGACTTAAAGCCGAAACCTGCTTACCTTACCTTACCTGGAAGAGAACT 540
Qy 542 CTAAGATCATGATCACTGACTTTGGTCTGTCCAGATGGAACAGATGCGATCATGTCGA 601
Db 541 CTAAGATCATGATCACTGACTTTGGTCTGTCCAGATGGAACAGATGCGATCATGTCGA 600
Qy 602 CTGCTGTGGGACCCAGGCTACGTGGCTCCAGAGTGTCTGGCCGACAGAACCCCTACAGCA 661
Db 601 CTGCTGTGGGACCCAGGCTACGTGGCTCCAGAGTGTCTGGCCGACAGAACCCCTACAGCA 660
Qy 662 AGGCTGTGGATGCTGTGGTCCATCGGCGTCACTACCTACATATTTGCTCTGTGGATACCCCC 721
Db 661 AGGCTGTGGATGCTGTGGTCCATCGGCGTCACTACCTACATATTTGCTCTGTGGATACCCCC 720
Qy 722 CGTTCTATGAGAAACCGAGTCTTAAGCTTTTCGAGAGATCAAGAGGCTACTATGAGT 781
Db 721 CATTTCTATGAGAAACCGAGTCTTAAGCTTTTCGAGAGATCAAGAGGCTACTATGAGT 780
```

```
Qy 782 TTGAGTCTCCATCTCTGGGATGACATTTCTCAGTCAGCCAGGACTTTATTTCGCCACTT 839
Db 781 TTGAGTCTCCATCTCTGGGATGACATTTCTCAGTCAGCCAGGACTTTATTTCGCCACTT 838

RESULT 2
BI821474
LOCUS      BI821474      809 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603038366F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179336 5',
mRNA sequence.
ACCESSION  BI821474
VERSION    BI821474.1 GI:15933024
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 809)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@email.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L2AM11447 row: c column: 17
            High quality sequence stop: 808.
FEATURES             Location/Qualifiers
     source            1..809
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5179336"
                     /clone_lib="NIH_MGC_115"
                     /lab_host="DH10B"
     note              "note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH-MGC Library."
BASE COUNT      212 a      192 c      218 g      187 t
ORIGIN
Query Match      31.1%; Score 762; DB 13; Length 809;
Best Local Similarity 100.0%; Pred. No. 5.1e-306;
Matches 762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGGAGTGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 60
Db 6 TGGAGTGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 65
Qy 61 CTGGAGGCAATGGGTGCGAAAGGAAGAGATGACTGCTGCTGGAAGAACAGACCACC 120
Db 66 CTGGAGGCAATGGGTGCGAAAGGAAGAGATGACTGCTGCTGGAAGAACAGACCACC 125
Qy 121 AACATCCGGAACCTTCATTTTATGGAAGTGTCTGGATCAGAGCTTCTCAGAAGTT 180
Db 126 AACATCCGGAACCTTCATTTTATGGAAGTGTCTGGATCAGAGCTTCTCAGAAGTT 185
Qy 181 TTCCTGGTGAAGCAAGACTGACTGGGAGCTCTTTCCTGGAAGTGCATCAAGAGTCA 240
Db 186 TTCCTGGTGAAGCAAGACTGACTGGGAGCTCTTTCCTGGAAGTGCATCAAGAGTCA 245
Qy 241 CCTGCTTCCGGGACAGCAGCTCGGAGATGAGATGCTGTGTGTGAAAGATCAAGCAT 300
```


|||||
Db CCTGCCCTCCGGACAGCAGCCTGGAGAAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 305
QY 301 GAAACATTTCTGACCTGGAGGACATCTATGAGAGACACACCACTACTACCTGGTTCATG 360
Db 306 GAAACATTTGACCTGGAGGACATCTATGAGAGACACCACTACTACCTGGTTCATG 365
QY 361 CAGCTTGTCTGTTGGGGAGCTCTTTGACCCGGATCCTGGAGCGGGTGTCTACACAGAG 420
Db 366 CAGCTTGTCTGTTGGGGAGCTCTTTGACCCGGATCCTGGAGCGGGTGTCTACACAGAG 425
QY 421 AAGGATCCAGTCTGGTGATCCAGCAGGCTTGTCCGCGAGTGAATACCTACATGAGAT 480
Db 426 AAGGATCCAGTCTGGTGATCCAGCAGGCTTGTCCGCGAGTGAATACCTACATGAGAT 485
QY 481 GGCATCTCCACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCTTACCTTACCTTACCT 540
Db 486 GGCATCTCCACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCTTACCTTACCTTACCT 545
QY 541 TCTAAGATCATGATGACTGCTTGTGCTGCTCAAGATGAACAGATGGCATCATGTC 600
Db 546 TCTAAGATCATGATGACTGCTTGTGCTGCTCAAGATGAACAGATGGCATCATGTC 605
QY 601 ACTGCCCTGGGACCCAGCTACGTGGCTCCAGAGTGTGCTGCCAGAACCTTACAGC 660
Db 606 ACTGCCCTGGGACCCAGCTACGTGGCTCCAGAGTGTGCTGCCAGAACCTTACAGC 665
QY 661 AAGGCTGTGATGCTGCTCCATCGGCTCATCACCTACATATGCTGCTGTGATACCC 720
Db 666 AAGGCTGTGATGCTGCTCCATCGGCTCATCACCTACATATGCTGCTGTGATACCC 725
QY 721 CCGTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATC 762
Db 726 CCGTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATC 767

RESULT 3
BI824483
LOCUS 60303885F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179957 5',
DEFINITION mRNA sequence.
ACCESSION BI824483
VERSION BI824483.1 GI:15936033
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LNA11448 row: m column: 14
High quality sequence stop: 709.
Location/Qualifiers
1. .740
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5179957"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1

male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 196 a 181 c 196 g 167 t
ORIGIN

Query Match 29.0%; Score 710; DB 13; Length 740;
Best Local Similarity 100.0%; Pred. No. 2.2e-284;
Matches 710; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGAGTGGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 60
Db 1 TGGAGTGGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 60
QY 61 CTGGAGGCAATGGGTGCGAAGGAAGATGACTGCAGTTCTTGGAAAGAACAGACCACC 120
Db 61 CTGGAGGCAATGGGTGCGAAGGAAGATGACTGCAGTTCTTGGAAAGAACAGACCACC 120
QY 121 AACATCCCGAAAAACCTTCATTTTATGGAAGTGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180
Db 121 AACATCCCGAAAAACCTTCATTTTATGGAAGTGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180
QY 181 TTCTCTGTGGAAGCAAGACTGACTGGGAAGCTTTTGTGCTCTGAAGTGCATCAAGAAGTCA 240
Db 181 TTCTCTGTGGAAGCAAGACTGACTGGGAAGCTTTTGTGCTCTGAAGTGCATCAAGAAGTCA 240
QY 241 CCTGCTTCCGGGACAGCAGCTGGAGATGAGATTGCTGTTCGAAAGATCAAGCAT 300
Db 241 CCTGCTTCCGGGACAGCAGCTGGAGATGAGATTGCTGTTCGAAAGATCAAGCAT 300
QY 301 GAAACATTTGTGACCTCGAGGACATCTATGAGAGCACCACCCTACTACTGCTCATG 360
Db 301 GAAACATTTGTGACCTCGAGGACATCTATGAGAGCACCACCCTACTACTGCTCATG 360
QY 361 CAGCTTGTTCGTGGGGAGCTTTTGACCGGATCTTGGAGCGGGTGTCTACACAGAG 420
Db 361 CAGCTTGTTCGTGGGGAGCTTTTGACCGGATCTTGGAGCGGGTGTCTACACAGAG 420
QY 421 AAGGATGCCAGTCTGCTGATCCAGCAGGCTTTGTGCGCAGTGAATACCTACATGAGAT 480
Db 421 AAGGATGCCAGTCTGCTGATCCAGCAGGCTTTGTGCGCAGTGAATACCTACATGAGAT 480
QY 481 GGCATGCTCCACAGAGACTTAAAGCCCGAAAAACCTTACCTTTACCTTACCTTACCTTACCT 540
Db 481 GGCATGCTCCACAGAGACTTAAAGCCCGAAAAACCTTACCTTTACCTTACCTTACCTTACCT 540
QY 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCAGAGATGGAACAAATGGCATCATGTCC 600
Db 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCCTCAAGATGGAACAAATGGCATCATGTCC 600
QY 601 ACTGCTGTGGGACCCAGGCTACGTGGCTCCAGAGTGTGCTGCCAGAACCCCTACAGAC 660
Db 601 ACTGCTGTGGGACCCAGGCTACGTGGCTCCAGAGTGTGCTGCCAGAACCCCTACAGAC 660
QY 661 AAGGCTGTGATGCTGTGCTCCATCGGCGTCACTACCATACATATGCTCTG 710
Db 661 AAGGCTGTGATGCTGTGCTCCATCGGCGTCACTACCATACATATGCTCTG 710

RESULT 4
BI084101
LOCUS 602869466F1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013856 5',
DEFINITION mRNA sequence.
ACCESSION BI084101
VERSION BI084101.1 GI:14502431
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1048)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

FEATURES	source
align quality sequence stop: 033. location/Qualifiers	
1. .1048	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMAGE:5013856"	
/clone_lib="NIH_MGC_102"	
/tissue_type="epidermoid carcinoma, cell line"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
231 a	298 c 264 g 254 t 1 others
BASE COUNT	

Query Match	28.0%;	Score	584;	DB	13;	Length	1048;
Best Local Similarity	100.0%;	Pred. No.	1.2e-273;				
Matches	684;	Conservative	0;	Mismatches	0;	Indels	0;
0;							
QY	1425	GGTCATGGTACCAAGTAAAGCAGTCAGGAGCTCCACACTCCCGGCAGGCGACACTGGAGT	1484				
DB	2	GGTCATGGTACCAAGTAAAGCAGTCAGGAGCTCCACACTCCCGGCAGGCGACACTGGAGT	61				
QY	1485	CTGCTCATTATGTCATCTCTGGAGCCTTGCCCTATGTCACCTGCAATTTTTCAGGAGACAT	1544				
DB	62	CTGCTCATTATGTCATCTCTGGAGCCTTGCCCTATGTCACCTGCAATTTTTCAGGAGACAT	121				
QY	1545	ATTCAACTCCTCTGCTCTTCCAAACCTCTGGTGTCTATCCGCAGAGGAGGAAGGCAGAGC	1604				
DB	122	ATTCAACTCCTCTGCTCTTCCAAACCTCTGGTGTCTATCCGCAGAGGAGGAAGGCAGAGC	181				
QY	1605	AAGTGGAGCAGGGCTTACGAGGAGCAGTCTTGCCAGAGAGCACAGCTGCTGCCAGGG	1664				
DB	182	AAGTGGAGCAGGGCTTACGAGGAGCAGTCTTGCCAGAGAGCACAGCTGCTGCCAGGG	241				
QY	1665	GGGCAGCCCTCATAGGAGGCCAGGAGGAGCCCCAAGCGGTAGAAGCGTTGTGTAAGC	1724				
DB	242	GGGCAGCCCTCATAGGAGGCCAGGAGGAGCCCCAAGCGGTAGAAGCGTTGTGTAAGC	301				
QY	1725	TGTGAGCAGGAGAAAGCGTGGCCACCAAGCTTCAGAGTCCCTGACCTGCCTGCTCTATG	1784				
DB	302	TGTGAGCAGGAGAAAGCGTGGCCACCAAGCTTCAGAGTCCCTGACCTGCCTGCTCTATG	361				
QY	1785	CCCCACACCCTACGTGCGCGTGGCTCTGTGACAGTGTACGTAGATAGCTCTCGCTCGGGTCT	1844				
DB	362	CCCCACACCCTACGTGCGCGTGGCTCTGTGACAGTGTACGTAGATAGCTCTCGCTCGGGTCT	421				
QY	1845	GTGCTGTTTGTGCGTAAAGCTTAATGGGCTGGCCAGGCTGTGCTACCTCTTCCAAAGCA	1904				
DB	422	GTGCTGTTTGTGCGTAAAGCTTAATGGGCTGGCCAGGCTGTGCTACCTCTTCCAAAGCA	481				
QY	1905	AGCCATATGGAGCATCTACCCAGCACTCCCACTCTGCACACACTCACTCCCACTCTCAAG	1964				

Db	482	AGCCATATGGAGCAATCTACCCAGAGTCCCACTCTGCACACACTCACTCCCACTCTCTCAAG	541
Qy	1965	CCTCAACCTCTTGGCCAGATTGGCTCATTAATGTCTGCTTGCCTGCCCATCTGCATGAT	2024
Db	542	CCTCAACCTCTTGGCCAGATTGGCTCATTAATGTCTGCTTGCCTGCCCATCTGCATGAT	601
Qy	2025	GACAGGCAGCTCCCAATGGTGTGCTGTGAGCTCTTCAAGTCTTAATCCTTAACTCC	2084
Db	602	GACAGGCAGCTCCCAATGGTGTGCTGTGAGCTCTTCAAGTCTTAATCCTTAACTCC	661
Qy	2085	AGGATTAGCTCCCAAGTGGCTGA	2108
Db	662	AGGATTAGCTCCCAAGTGGCTGA	685
RESULT	5		
LOCUS	BM807335	1129 bp	linear EST 05-MAR-2002
DEFINITION	AGENCOURT_6575096 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5732298		
	5', mRNA sequence.		
ACCESSION	BM807335		
VERSION	BM807335.1	GI:19124158	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1129)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: Invitrogen		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	Plate: LLAM12734 row: c column: 19		
	High quality sequence start: 11		
	High quality sequence stop: 651.		
FEATURES	Location/Qualifiers		
source	1..1129		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5732298"		
	/clone_lib="NIH_MGC_124"		
	/tissue_type="hippocampus"		
	/lab_host="DH10B"		
	/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV (destroyed); Site_2: NotI; RNA source male hippocampus, age 27. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."		
BASE COUNT	287 a 308 c 294 g 240 t		
ORIGIN			
Query Match	27.2%	Score 665;	DB 14; Length 1129;
Best Local Similarity	100.0%;	Pred. No. 9.5e-266;	
Matches	665; Conservative	0; Mismatches	0; Indels
			0; Gaps
			0;
Qy	1	TGGAGTGGGAGCTCAACGAGGATCTTCCCGAGTCCCTGGCATCTCTCAGAGCTTCAACT	60
Db	36	TGGAGTGGGAGCTCAACGAGGATCTTCCCGAGTCCCTGGCATCTCTCAGAGCTTCAACT	95
Qy	61	CTGAGGCAATGGTCTGAAAGAGAGATGACTGCGAGTTCCTTGGAAAGAAACAGACCAC	120

Db 96 CTGGAGCAATGGTTCGAAAGGAAGATGACTCAGTTCCTGCGAAGAAACAGACCACC 155
QY 121 AACATCCGGAACACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTTCTCAGAACTT 180
Db 156 AACATCCGGAACACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTTCTCAGAACTT 215
QY 181 TTCTCTGGTCAAGCAAGACTGACTGGGAAGCTTTTGTCTCTGAAAGTGCATCAAGAAGTCA 240
Db 216 TTCTCTGGTCAAGCAAGACTGACTGGGAAGCTTTTGTCTCTGAAAGTGCATCAAGAAGTCA 275
QY 241 CCTGCTTCCGGGACAGCAGCTTGGGAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 300
Db 276 CCTGCTTCCGGGACAGCAGCTTGGGAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 335
QY 301 GAAACATTTGTGACCTGGAGACATCTATGAGAGCACACCCTACTACTGCTGTCATG 360
Db 336 GAAACATTTGTGACCTGGAGACATCTATGAGAGCACACCCTACTACTGCTGTCATG 395
QY 361 CAGCTTGTCTTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAG 420
Db 396 CAGCTTGTCTTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAG 455
QY 421 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTTGTCGCGAGTGAATACTACATGAGAAT 480
Db 456 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTTGTCGCGAGTGAATACTACATGAGAAT 515
QY 481 GGCATCGTCCAGAGACTTAAGCCCGAAGACCTTGCTTTACCTTACCCCTGAAGAGAAC 540
Db 516 GGCATCGTCCAGAGACTTAAGCCCGAAGACCTTGCTTTACCTTACCCCTGAAGAGAAC 575
QY 541 TCTAAGATCATGACTGACTTGTGTCGCAAGTGGACACAGAAATGGGATCATGTC 600
Db 576 TCTAAGATCATGACTGACTTGTGTCGCAAGTGGACACAGAAATGGGATCATGTC 635
QY 601 ACTGCTGTGGGACCCAGGCTACGTGGTCCAGAAAGTCTGGGCCAGAAACCCCTACAGC 660
Db 636 ACTGCTGTGGGACCCAGGCTACGTGGTCCAGAAAGTCTGGGCCAGAAACCCCTACAGC 695
QY 661 AAGGC 665
Db 696 AAGGC 700

RESULT 6
BI084897/c
LOCUS 602869466r1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013856 3',
DEFINITION mRNA sequence.
ACCESSION BI084897
VERSION BI084897.1 GI:14503227
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 966)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1820 row: d column: 17
High quality sequence stop: 832.
Location/Qualifiers
1..966
/organism="Homo sapiens"

FEATURES
source

/db_xref="taxon:9606"
/clone="IMAGE:5013856"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: salivary gland; Vector: pOTB7; Site_1: xhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 219 a 238 c 302 g 207 t
ORIGIN
Query Match 26.8%; Score 656; DB 13; Length 966;
Best Local Similarity 99.9%; Pred. No. 5.5e-262;
Matches 776; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1589 GGGAGGAGGAGCAGAGTGGAGCAGGGCTTAGCAGAGCAGAGTTTCTGGCCAGAGCAC 1648
Db 797 GGGAGGAGGAGCAGAGTGGAGCAGGGCTTAGCAGAGCAGAGTTTCTGGCCAGAGCAC 738
QY 1649 CAGCCTGCTGCCAGCGGGCAGCCCTCATAGGAGGCCAGGAGGAGGCCCAAGGCGTA 1708
Db 737 CAGCCTGCTGCCAGCGGGCAGCCCTCATAGGAGGCCAGGAGGAGGCCCAAGGCGTA 678
QY 1709 GAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACAGCTTCCAGAGTCTCCCTG 1768
Db 677 GAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACAGCTTCCAGGTCCTCCTG 618
QY 1769 ACCTGCTGCTCATATGCCACACCCCTACGTGCCGTGGCTCTGTGCAAGTGTACGTAGATA 1828
Db 617 ACCTGCTGCTCATATGCCACACCCCTACGTGCCGTGGCTCTGTGCAAGTGTACGTAGATA 558
QY 1829 GCTCTGCGCTGGGTCGTGCTGTTGTCGTAAGAGCTTAATGGCGGCCAGGCTCTGT 1888
Db 557 GCTCTGCGCTGGGTCGTGCTGTTGTCGTAAGAGCTTAATGGCGGCCAGGCTCTGT 498
QY 1889 CACCTTCTCCAAAGCAAGCCATATGAGAGCTTACCAGAGCTCCCACTCTGCACACACTC 1948
Db 497 CACCTTCTCCAAAGCAAGCCATATGAGAGCTTACCAGAGCTCCCACTCTGCACACACTC 438
QY 1949 ACTCCACCTCTCAAGCCTCCACCTCTTGCCAGAGTTGGCTCATTAATGTCGTCCT 2008
Db 437 ACTCCACCTCTCAAGCCTCCACCTCTTGCCAGAGTTGGCTCATTAATGTCGTCCT 378
QY 2009 GCCCATCTGCATGAATGACAGGAGCTCCCATGCTGCTGCTGCTGAGCTCTTCAAGT 2068
Db 377 GCCCATCTGCATGAATGACAGGAGCTCCCATGCTGCTGCTGAGCTCTTCAAGT 318
QY 2069 TCTAATCCTTAACCTCCAGGATTAGCTCCCAAGTGCCTGAGACCCAGCCAGCAGCTTCT 2128
Db 317 TCTAATCCTTAACCTCCAGGATTAGCTCCCAAGTGCCTGAGACCCAGCCAGCAGCTTCT 258
QY 2129 GGCCCTTCTCCTCGCTCAATPCTAAAAGCAGTGCACACCCCTCCAAAGTGAATAGAAAG 2188
Db 257 GGCCCTTCTCCTCGCTCAATPCTAAAAGCAGTGCACACCCCTCCAAAGTGAATAGAAAG 199
QY 2189 AAGTTCATGAGTAAAGGCTGCAAGGAATTTCTATCTTGGCCACATGTCCTCCGTGCACAC 2248
Db 198 AAGTTCATGAGTAAAGGCTGCAAGGAATTTCTATCTTGGCCACATGTCCTCCGTGCACAC 139
QY 2249 ACCCAATGGAGTTAACTTGAAGTTGACTATTTAATGCTGCGCAGAGTTCTTAATCCT 2308
Db 138 ACCCAATGGAGTTAACTTGAAGTTGACTATTTAATGCTGCGCAGAGTTCTTAATCCT 79
QY 2309 GCCTCTGTTCCTTTTCTCTCTTTGAAAGTCCAGCAGCAGCTTCTGTCTTCCCTCCCA 2365
Db 78 GCCTCTGTTCCTTTTCTCTCTTTGAAAGTCCAGCAGCAGCTTCTGTCTTCCCTCCCA 22

```
RESULT 7
BI753035      731 bp      mRNA      linear      EST 25-SEP-2001
LOCUS         603025844F1 NIH_MGC_114 Homo sapiens cdna clone IMAGE:5196323 5',
DEFINITION   mRNA sequence.
ACCESSION    BI753035
VERSION      BI753035.1 GI:15744613
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 731)
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE       NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL     National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT     Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgaaps-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLNL1491 row: g column: 12
            High quality sequence stop: 729.
FEATURES     Location/Qualifiers
             1. 731
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5196323"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
                Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
                male brains, age range 23-27 yo. Library is oligo-dT
                primed and directionally cloned (EcoRV site is destroyed
                upon cloning). Average insert size 1.5 kb, insert size
                range 1-3 kb. Library is normalized and enriched for
                full-length clones and was constructed by C. Gruber
                (Invitrogen). Research Genetics tracking code 019. Note:
                this is a NIH_MGC Library."
BASE COUNT   190 a 182 c 191 g 168 t
ORIGIN
Query Match      26.6%; Score 650; DB 13; Length 731;
Best Local Similarity 100.0%; Pred. No. 1.9e-259;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 73 GGTCAAGGAAGAGATGACTCGATCTCTGGGAGAACACACACATCCGGAAA 132
DB 73 GGTCAAGGAAGAGATGACTCGATCTCTGGGAGAACACACACATCCGGAAA 132
QY 133 ACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAGTTTCTCGTGAAG 192
DB 133 ACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAGTTTCTCGTGAAG 192
QY 193 CAAGACTGACTGGGAAGCTTTTGTCTGAAGTCATCAAGAAAGTCACCTGCCCTCCGG 252
DB 193 CAAGACTGACTGGGAAGCTTTTGTCTGAAGTCATCAAGAAAGTCACCTGCCCTCCGG 252
QY 253 GACAGCAGCTGGAGATGAGATGCTGTGTGAAAGATCAAGCATGAACATTTGTG 312
DB 253 GACAGCAGCTGGAGATGAGATGCTGTGTGAAAGATCAAGCATGAACATTTGTG 312
QY 313 ACCCTGGAGGACATCTATGAGACACACACCTACTACCTGGTCATCCAGCTTTTCT 372
DB 313 ACCCTGGAGGACATCTATGAGACACACACCTACTACCTGGTCATCCAGCTTTTCT 372
QY 373 GGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATCCAGT 432
DB 373 GGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATCCAGT 432
```

```
QY 433 CTGGTGATCCAGCAGGTCTTGTGCGCAGTGAATAACCTACATGAGATGSCATCGTCCAC 492
DB 433 CTGGTGATCCAGCAGGTCTTGTGCGCAGTGAATAACCTACATGAGATGSCATCGTCCAC 492
QY 493 AGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTTAAGATCATG 552
DB 493 AGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTTAAGATCATG 552
QY 553 ATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCCTGCTGCTGGG 612
DB 553 ATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCCTGCTGCTGGG 612
QY 613 ACCCCAGGCTACGTGGCTCCAGAAAGTCTGGCCCGAGAAACCCCTACAGCAAGGCTGTGGAT 672
DB 613 ACCCCAGGCTACGTGGCTCCAGAAAGTCTGGCCCGAGAAACCCCTACAGCAAGGCTGTGGAT 672
QY 673 TGCTGGTCCATCGGCTCATCACTACATATTCGCTCTGTGGATACCCCCC 722
DB 673 TGCTGGTCCATCGGCTCATCACTACATATTCGCTCTGTGGATACCCCCC 722

RESULT 8
BI818261      742 bp      mRNA      linear      EST 04-OCT-2001
LOCUS         603032510F1 NIH_MGC_115 Homo sapiens cdna clone IMAGE:5173587 5',
DEFINITION   mRNA sequence.
ACCESSION    BI818261
VERSION      BI818261.1 GI:15928724
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 742)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgaaps-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLNL1432 row: d column: 04
            High quality sequence stop: 742.
FEATURES     Location/Qualifiers
             1. 742
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5173587"
                /lab_host="DH10B"
                /note="Organ: pooled brain, lung, testis; Vector:
                pCMV-SPORT6; Site: 1: NotI; Site 2: EcoRV (destroyed); RNA
                source anonymous pool of 6 male brains, age range 23-27; 1
                male lung, age 27; and 1 male testis, age 69. Library is
                oligo-dT primed and directionally cloned (EcoRV site is
                destroyed upon cloning). Average insert size 1.8 kb,
                insert size range 1-3 kb. Library is normalized and
                enriched for full-length clones and was constructed by C.
                Gruber (Invitrogen). Research Genetics tracking code
                021. Note: this is a NIH_MGC Library."
BASE COUNT   197 a 180 c 197 g 168 t
ORIGIN
Query Match      26.0%; Score 636; DB 13; Length 742;
Best Local Similarity 99.9%; Pred. No. 1.2e-253;
Matches 686; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 17 GCAGGATCTTCCGAGTCCCTGGCATCCTCAGAGCTTCAACTCTGGAGCAATGGGTC 76
Db 24 GCAGGATCTTCCGAGTCCCTGGCATCCTCAGAGCTTCAACTCTGGAGCAATGGGTC 83
QY 77 GAAAGGAAGAAGATGACTGCAGTTCCTGGAAGAAACAGACCACCAACATCCGGAACCT 136
Db 84 GAAAGGAAGAAGATGACTGCAGTTCCTGGAAGAAACAGACCACCAACATCCGGAACCT 143
QY 137 TCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAAGTCTTCTGGTGAAGCAAA 196
Db 144 TCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAAGTCTTCTGGTGAAGCAAA 203
QY 197 GACTGACTGGGAAGCTTTTGTCTGAAGTGCATCAAGAGTCACTGCCTTCCGGGACA 256
Db 204 GACTGACTGGGAAGCTTTTGTCTGAAGTGCATCAAGAGTCACTGCCTTCCGGGACA 263
QY 257 GCAGCTGGAGATGAGATGCTGTGTGTTGAAAAGATCAAGCATGAAACATTTGTGACCC 316
Db 264 GCAGCTGGAGATGAGATGCTGTGTGTTGAAAAGATCAAGCATGAAACATTTGTGACCC 323
QY 317 TGGAGACATCTATGAGAGCACCACCTACTACTGTCATGTCAGCTTGTCTGGTG 376
Db 324 TGGAGACATCTATGAGAGCACCACCTACTACTGTCATGTCAGCTTGTCTGGTG 383
QY 377 GGGAGCTCTTTGACCGATCTTGGAGCGGGTGTCTACACAGAGAGATGCCAGTCTGG 436
Db 384 GGGAGCTCTTTGACCGATCTTGGAGCGGGTGTCTACACAGAGAGATGCCAGTCTGG 443
QY 437 TGATCCAGCAGGTCTTTGTCGGCAGTGAATACCTACATGAGAAATGGCATCTGCCACAG 496
Db 444 TGATCCAGCAGGTCTTTGTCGGCAGTGAATACCTACATGAGAAATGGCATCTGCCACAG 503
QY 497 ACTTAAGCCCGAAACCTTGTACCTTACCCCTGAGAGAACTCTTAGATCATGATCA 556
Db 504 ACTTAAGCCCGAAACCTTGTACCTTACCCCTGAGAGAACTCTTAGATCATGATCA 563
QY 557 CTGACTTTGCTGTCTCCAGATGGAACAGATGGCATATGCCATCTGCCTGTGGGACCC 616
Db 564 CTGACTTTGCTGTCTCCAGATGGAACAGATGGCATATGCCATCTGCCTGTGGGACCC 623
QY 617 CAGGCTACGTGGCTCCAGAGTGTCTGGCCAGAGAAACCTTACAGCAAGGCTGTGGATTGCT 676
Db 624 CAGGCTACGTGGCTCCAGAGTGTCTGGCCAGAGAAACCTTACAGCAAGGCTGTGGATTGCT 683
QY 677 GFTCCATCGGCTCATCCTACATAT 703
Db 684 GFTCCATCGGCTCATCCTACATAT 710

RESULT 9
BI772626
LOCUS 812 bp mRNA linear EST 25-SEP-2001
DEFINITION 603060879F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5210299 5',
mRNA sequence.
ACCESSION BI772626
VERSION BI772626.1 GI:15764204
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 812)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LUAMI1527 row: m column: 20
High quality sequence start: 2
High quality sequence stop: 808.
Location/Qualifiers
source 1..812
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5210299"
/lab_lib="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site:1: NotI; Site:2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed, and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 217 a 186 c 221 g 188 t
ORIGIN

Query Match 23.6%; Score 577; DB 13; Length 812;
Best Local Similarity 99.8%; Pred. No. 3.9e-229;
Matches 627; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TGGAGTGGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 60
Db 9 TGGAGTGGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 68
QY 61 CTGAGGCAATGGGTGCAAGAGAGAGATGACTGCTGAGTTCTCTGGAGAAACAGACACC 120
Db 69 CTGAGGCAATGGGTGCAAGAGAGAGATGACTGCTGAGTTCTCTGGAGAAACAGACACC 128
QY 121 AACATCCGGAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180
Db 129 AACATCCGGAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 188
QY 181 TTCCTGCTGGAAGCAAGACTGACTGGGAAGCTCTTCTCTGAAAGTGCATCAAGAAAGTCA 240
Db 189 TTCCTGCTGGAAGCAAGACTGACTGGGAAGCTCTTCTCTGAAAGTGCATCAAGAAAGTCA 248
QY 241 CCTGCCTTCGGGACAGCAGCTGGAGAAATGAGATTGCTGTGTGAAAAGATCAAGCAT 300
Db 249 CCTGCCTTCGGGACAGCAGCTGGAGAAATGAGATTGCTGTGTGAAAAGATCAAGCAT 308
QY 301 GAAACATTTGACCCCTGGAGGACATCTATGAGAGCACCACCACCTACTACCTGGTCATG 360
Db 309 GAAACATTTGACCCCTGGAGGACATCTATGAGAGCACCACCACCTACTACCTGGTCATG 368
QY 361 CAGCTTCTTCTGTTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAG 420
Db 369 CAGCTTCTTCTGTTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAG 428
QY 421 AAGGATCCAGTCTGGTGTATCCAGCAGGCTTGTTCGCGCATGAAATACCTACATGAGAT 480
Db 429 AAGGATCCAGTCTGGTGTATCCAGCAGGCTTGTTCGCGCATGAAATACCTACATGAGAT 488
QY 481 GGCATCTGCACAGAGACTTTAAAGCCGAAAACCTGCTTTACCTTACCCCTGAAGAGAAC 540
Db 489 GGCATCTGCACAGAGACTTTAAAGCCGAAAACCTGCTTTACCTTACCCCTGAAGAGAAC 548
QY 541 TCTAAGATCATGATFACCTGACTTTGTTGCTCTCCAAAGATGGAACAGAAATGGCATCATGTCC 600
Db 549 TCTAAGATCATGATFACCTGACTTTGGGTCTGTCCAAGATGGAACAGAAATGGCATCATGTCC 608
QY 601 ACTGCTCTGGGACCCAGGCTACGTGG 628
Db 609 ACTGCTCTGGGACCCAGGCTACGTGG 636

```

RESULT 10
BM921532      1068 bp      mRNA      linear      EST 12-MAR-2002
LOCUS      AGENCOURT_6708041 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753010
DEFINITION      5', mRNA sequence.
ACCESSION      BM921532
VERSION      BM921532.1 GI:19371911
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1068)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12788 row: b column: 19
High quality sequence stop: 592.
FEATURES             source
     1..1068
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="IMAGE:5753010"
     /clone_lib="NIH_MGC_115"
     /lab_host="DH10B"
     /note="Organ: pooled brain, lung, testis; Vector:
pcMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT      279 a      281 c      260 g      248 t
ORIGIN
Query Match      22.5%; Score 551; DB 14; Length 1068;
Best Local Similarity 99.8%; Pred. No. 2.3e-218;
Matches 601; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1      TGGAGTGGGAGCTCAAGCAGGATTCTCCCGAGTCCTCGCATCCTCAGAACGTTCAACT 60
|||||
Db      14      TGGAGTGGGAGCTCAAGCAGGATTCTCCCGAGTCCTCGCATCCTCAGAACGTTCAACT 73
|||||
QY      61      CTGGAGCAATGGTGCAGGAAGAGATGACTGCAGTTCCTGGAAGAAACAGACCACC 120
|||||
Db      74      CTGGAGCAATGGTGCAGGAAGAGATGACTGCAGTTCCTGGAAGAAACAGACCACC 133
|||||
QY      121     AACATCGGAAACCTTCATTTTATGGAAGTCCTGGGATCAGGAGCTTTCAGAGATT 180
|||||
Db      134     AACATCGGAAACCTTCATTTTATGGAAGTCCTGGGATCAGGAGCTTTCAGAGATT 193
|||||
QY      181     TTCTGTGTGAAGCAAAAGACTGACTGGGAAGCTTTTGCTCTGAAGTCATCAAGAAGTCA 240
|||||
Db      194     TTCTGTGTGAAGCAAAAGACTGACTGGGAAGCTTTTGCTCTGAAGTCATCAAGAAGTCA 253
|||||
QY      241     CCTGCCTCCGGACAGCAGCCCTGGAGATGAGATTGCTGTGTGTTGAAAAGATCAAGCAT 300
|||||
Db      254     CCTGCCTCCGGACAGCAGCCCTGGAGATGAGATTGCTGTGTGTTGAAAAGATCAAGCAT 313
|||||
QY      301     GAAAAATGTGACCTGGAGGACATCTATGAGAGCACCACCCTACTACCTGGTGCATG 360
|||||

```

```

Db      314     GAAAAATGTGACCTGGAGGACATCTATGAGAGCACCACCCTACTACCTGGTGCATG 373
QY      361     CAGCTTGTTCCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAG 420
|||||
Db      374     CAGCTTGTTCCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACACAGAG 433
|||||
QY      421     AAGATGCCAGTCTGGTGATCCAGCAGGTCTTGTCCGAGTGAATACCTACATGAGAAT 480
|||||
Db      434     AACGATGCCAGTCTGGTGATCCAGCAGGTCTTGTCCGAGTGAATACCTACATGAGAAT 493
|||||
QY      481     GGCATCTGCACACAGACTTAAAGCCCGAAACCTGCTTTACCTTACCTCCCTGAAGAGAAC 540
|||||
Db      494     GGCATCTGCACACAGACTTAAAGCCCGAAACCTGCTTTACCTTACCTCCCTGAAGAGAAC 553
|||||
QY      541     TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTC 600
|||||
Db      554     TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTC 613
|||||
QY      601     AC 602
Db      614     AC 615

RESULT 11
BQ102407/c
LOCUS      BQ102407
DEFINITION      BQ102407.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:6135181 3', mRNA sequence.
ACCESSION      BQ102407
VERSION      BQ102407.1 GI:20135391
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 479)
AUTHORS      Melton D., Brown J., Kenty G., Permutt, A., Lee C., Kaestner, K.,
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wyllie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other_ESTs:      ij2ic07.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center This clone is
available royalty-free through LLNL; please contact the IMAGE
consortium (info@image.llnl.gov) for further information
Seq primer: -40UP from Gibco
High quality sequence stop: 448.
FEATURES             source
     1..479
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="IMAGE:6135181"
     /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
     /sex="Both"
     /tissue_type="Islets of Langerhans"
     /dev_stage="Adult"
     /lab_host="DH10B"
     /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using
SuperScript plasmid library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column

```

fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 136 a 89 c 147 g 107 t

Query Match 19.6%; Score 479; DB 14; Length 479;

Best Local Similarity 100.0%; Pred. No. 2.5e-188; Length 479;

Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1960 TCAAGCCTCAACCTCTTGGCCAGATTGGGCTCATTAAATGTGCTGCTGCCCATCTGCA 2019

Db 479 TCAAGCCTCAACCTCTTGGCCAGATTGGGCTCATTAAATGTGCTGCTGCCCATCTGCA 420

QY 2020 TGAATGACAGGACGCTCCCATGCTGCTGCTGAGCTCTTCAAGTTCTAAATCCTTA 2079

Db 419 TGAATGACAGGACGCTCCCATGCTGCTGCTGAGCTCTTCAAGTTCTAAATCCTTA 360

QY 2080 ACTCCAGGATTAGCTCCCAAGTSCGCTGAGACCCAGCCAGCACACTTCTGCCCTTCTCC 2139

Db 359 ACTCCAGGATTAGCTCCCAAGTSCGCTGAGACCCAGCCAGCACACTTCTGCCCTTCTCC 300

QY 2140 CTCGCTCAATCTAAAAGCAGTGCACACACCTCCAAAGTGGAAATAGAAAGTTCATGAG 2199

Db 299 CTCGCTCAATCTAAAAGCAGTGCACACACCTCCAAAGTGGAAATAGAAAGTTCATGAG 240

QY 2200 TAAGGGCTGCAAGGAATCTTATCTGCGCCACATGCTCTCCGTGCACACACCAATGGAG 2259

Db 239 TAAGGGCTGCAAGGAATCTTATCTGCGCCACATGCTCTCCGTGCACACACCAATGGAG 180

QY 2260 TTAACCTTGAAGTTCACATATTTTAAATGCTGCCAGGAGTTCTAATCTGCTGCTTCC 2319

Db 179 TTAACCTTGAAGTTCACATATTTTAAATGCTGCCAGGAGTTCTAATCTGCTGCTTCC 120

QY 2320 CTTTCTCTCTTGAAGTCCAGCACACCATCTTGTCTCTCCCGAGTTTCTCGCCCTC 2379

Db 119 CTTTCTCTCTTGAAGTCCAGCACACCATCTTGTCTCTCCCGAGTTTCTCGCCCTC 60

QY 2380 CACCCCTCAGCTTCAATGCTCAGTGTGCTTAAATAAATGGACATATTTTCTCTAA 2438

Db 59 CACCCCTCAGCTTCAATGCTCAGTGTGCTTAAATAAATGGACATATTTTCTCTAA 1

RESULT 12

LOCUS BG715920

DEFINITION BG715920 824 bp mRNA linear EST 08-MAY-2001

602676667F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799224 5',

mRNA sequence.

ACCESSION BG715920

VERSION BG715920.1 GI:13995107

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 824)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LIAM0688 row: e column: 17

High quality sequence stop: 766.

Location/Qualifiers

source 1. .824

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4799224"

/clone_lib="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag

); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTNN-3',

size-selected for average insert size 2.3 kb and

normalized to R0F 5. This is a primary library enriched

for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIH/NHGRI, National

Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 223 a 194 c 219 g 188 t

ORIGIN

Query Match 19.3%; Score 473; DB 12; Length 824;

Best Local Similarity 99.5%; Pred. No. 6.6e-186;

Matches 623; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGGAGTGGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 60

Db 38 TGGAGTGGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 97

QY 61 CTGGAGGCAATGGGTGCGAAGGAAAGAGATGACTGTCAGTTCTTCTGGAAGAAACAGACACC 120

Db 98 CTGGAGGCAATGGGTGCGAAGGAAAGAGATGACTGTCAGTTCTTCTGGAAGAAACAGACACC 157

QY 121 AACATCCCGAAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGATT 180

Db 158 AACATCCCGAAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGATT 217

QY 181 TTCTGCTGGAAGCAAGAACTGACTGGGAAGCTCTTTGCTCTGAAAGTGCATCAAGAAAGTCA 240

Db 218 TTCTGCTGGAAGCAAGAACTGACTGGGAAGCTCTTTGCTCTGAAAGTGCATCAAGAAAGTCA 277

QY 241 CCTGCTTCCGGGACAGCAGCCTTGAGAAATGAGATTGCTGTGTTGAAAGATCAAGCAT 300

Db 278 CCTGCTTCCGGGACAGCAGCCTTGAGAAATGAGATTGCTGTGTTGAAAGATCAAGCAT 337

QY 301 GAAACATTTGACCCCTGGAGGACATCTATGAGAGCACCCACCTACTACTCTGTCATG 360

Db 338 GAAACATTTGACCCCTGGAGGACATCTATGAGAGCACCCACCTACTACTCTGTCATG 397

QY 361 CAGCTTGTCTTCTGTTGGGAGCTCTTTGACCGGATCTCTGGAGCGGGTGTCTACACAGAG 420

Db 398 CAGCTTGTCTTCTGTTGGGAGCTCTTTGACCGGATCTCTGGAGCGGGTGTCTACACAGAG 457

QY 421 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTTCGGCAGTGAATACCTACATGAGAAT 480

Db 458 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTTCGGCAGTGAATACCTACATGAGAAT 517

QY 481 GGCATCTCCACAGAGACTTAAAGCCCGAAAACCTGCTTTACCTTACCCCTGAGAGAAC 540

Db 518 GGCATCTCCACAGAGACTTAAAGCCCGAAAACCTGCTTTACCTTACCCCTGAGAGAAC 577

QY 541 TCTAAGATCATGATCAGTCTGCTTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCC 600

Db 578 TCTAAGATCATGATCAGTCTGCTTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCC 637

QY 601 ACTGCTGTGGGACCCCAAGGCTACGT 626

Db 638 ACTGCTGTGGGACCCCAAGGCTACGT 663

BASE COUNT	ORIGIN
134 a	89 c 142 g 119 t
<p>NCI_CGAP_Col0 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE Cloneds 1057416-1061255, 114384-1145351) The resulting subtracted library contained 4 million recombinants. Subtraction was performed as previously described [Bonaldo , Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 5, 791-806. TAG_L1B-NCI_CGAP_Brn26 TAG_T1SSUE-brain TAG_SEQ=ATAGG"</p>	

Query Match	17.48;	Score 426;	DB 10;	Length 484;
Best Local Similarity	100.0%;	Pred. No. 2.6e-166;		
Matches 426;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2013	ATCTGCATGAATGACAGGCAGCTCCCCATGCTGGTCTGCTGTGAGCTCTTCAAGTTCTA	2072	
Db				
QY	2073	ATCCTTAACTCAGAGTTAGCTCCCAAGTGGCTGTGAGACCCAGCCAGCACACATTCCTGGCC	2132	
Db				
QY	381	ATCCTTAACTCCAGAGTTAGCTCCCAAGTGGCTGTGAGACCCAGCCAGCACATTCCTGGCC	322	
Db				
QY	2133	CTTCTCCCTGCCCTCAATCTTAAAGCAGTGGCCACACCCCTCCAAAGTGGAAATAGAAAGAAGT	2192	
Db				
QY	2193	TCATGAGTAAGGGCTGCAAGGAATCTTTATCTCTGGGCCACATGTCCTCGTGCACACACC	2252	
Db				
QY	2253	AATGGAGTTAACTTTGGAGTTGACTATTTTAAATGCTGCAGAGTTCCTAATCCTGGCCT	2312	
Db				
QY	201	AATGGAGTTAACTTTGGAGTTGACTATTTTAAATGCTGCAGAGTTCCTAATCCTGGCCT	142	
Db				
QY	2313	CTGTTCCCTTTTCTCTCCTTGAAGTCCAGCACACCATTCTTCTCTCTCCCGAGTTTCT	2372	
Db				
QY	2373	CGCCCTCCACCCCTCCAGCTTCAGCTCAGTGGTTGCTCTTAATAAATGACATATTTT	2432	
Db				
QY	2433	CTCTAA 2438		
Db				
QY	21	CTCTAA 16		
Db				

RESULT 15	AA838372/c	AA838372	410 bp	mRNA	linear	EST 09-JUN-1998			
LOCUS	OF29a11.s1	NCI_CGAP_Kid6	Homo sapiens	cdna	clone	IMAGE:1422524,			
DEFINITION	mRNA sequence.								
ACCESSION	AA838372	AA838372.1	GI:2913171						
VERSION	EST.								
KEYWORDS	human,								
SOURCE	Homo sapiens								
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
REFERENCE	1 (bases 1 to 410)								
AUTHORS	NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap .							
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index								
JOURNAL	Unpublished (1997)								
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@aim.nih.gov Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D. cdna Library Preparation: Stratagene, Inc. cdna Library Arrayed by: Greg Lennon, Ph.D.								

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 17:25:25 ; Search time 97 Seconds
(without alignments)
7736.475 Million cell updates/sec

Title: US-09-960-643-1
Perfect score: 2447
Sequence: 1 tggagtggagctcaagcag.....ttttctctataaaaaaaa 2447

Scoring table: OLIGO.NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCFUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	287	11.7	362	4 US-08-905-223-89	Sequence 89, Appl
2	23	0.9	296	1 US-08-738-367-8	Sequence 8, Appl
3	23	0.9	3380	2 US-09-156-425-1	Sequence 1, Appl
4	22	0.9	755	2 US-08-737-980-1	Sequence 1, Appl
5	21	0.9	1282	2 US-08-878-989-12	Sequence 12, Appl
6	21	0.9	1282	4 US-09-272-796-12	Sequence 12, Appl
7	20	0.8	3116	1 US-08-149-103-2	Sequence 2, Appl
8	20	0.8	3116	1 US-08-451-883-2	Sequence 2, Appl
9	20	0.8	3307	1 US-08-456-647B-5	Sequence 5, Appl
10	20	0.8	3307	2 US-08-237-401A-5	Sequence 5, Appl
11	20	0.8	3785	1 US-08-445-640-9	Sequence 9, Appl
12	20	0.8	3785	3 US-08-170-558-9	Sequence 9, Appl
13	20	0.8	3785	3 US-08-447-314-9	Sequence 9, Appl
14	20	0.8	3785	3 US-08-445-461-9	Sequence 9, Appl
15	19	0.8	1649	2 US-08-466-120-1	Sequence 1, Appl
16	19	0.8	1649	5 PCT-US94-07266-1	Sequence 1, Appl
17	19	0.8	2025	4 US-09-247-155-58	Sequence 58, Appl
18	19	0.8	2573	2 US-08-884-681-2	Sequence 2, Appl
19	19	0.8	2573	4 US-09-258-643-2	Sequence 2, Appl
20	19	0.8	2746	2 US-09-016-000-12	Sequence 12, Appl
21	19	0.8	55827	4 US-09-813-133A-3	Sequence 3, Appl
22	18	0.7	326	1 US-08-700-575-3	Sequence 3, Appl
23	18	0.7	973	3 US-09-013-881-16	Sequence 16, Appl
24	18	0.7	1001	4 US-09-641-638-444	Sequence 444, App
25	18	0.7	1001	4 US-09-641-638-446	Sequence 446, App
26	18	0.7	1283	4 US-09-282-305-11	Sequence 11, Appl
27	18	0.7	1631	4 US-09-051-239A-1	Sequence 1, Appl

c	28	18	0.7	1665	4	US-09-247-155-72	Sequence 72, Appl
	29	18	0.7	1813	4	US-08-123-934A-1	Sequence 1, Appl
	30	18	0.7	1813	5	PCT-US94-10080-1	Sequence 1, Appl
	31	18	0.7	2056	4	US-08-158-735A-1	Sequence 13, Appl
	32	18	0.7	2070	4	US-09-382-256-13	Sequence 13, Appl
	33	18	0.7	2070	4	US-09-395-115-13	Sequence 13, Appl
	34	18	0.7	2070	4	US-08-436-265-13	Sequence 13, Appl
	35	18	0.7	2070	4	US-09-679-187-13	Sequence 11, Appl
	36	18	0.7	2402	4	US-08-462-467B-11	Sequence 3, Appl
	37	18	0.7	2402	4	US-08-156-735A-3	Sequence 11, Appl
c	38	18	0.7	2402	4	US-08-334-179A-11	Sequence 1, Appl
	39	18	0.7	2846	4	US-08-915-795-1	Sequence 55, Appl
	40	18	0.7	5538	2	US-08-231-193A-55	Sequence 55, Appl
	41	18	0.7	5538	3	US-08-486-273A-55	Sequence 55, Appl
	42	18	0.7	5538	4	US-08-940-086A-55	Sequence 55, Appl
	43	18	0.7	5538	4	US-08-940-035A-55	Sequence 55, Appl
	44	18	0.7	5538	4	US-08-935-105A-55	Sequence 55, Appl
	45	18	0.7	5538	4	US-09-648-797-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-905-223-89
; Sequence 89, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Americ
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig.peptide
; LOCATION: 87..191
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 5.6
; OTHER INFORMATION: seq FIFMEVLGSGAFS/EV
US-08-905-223-89

```
Query Match      11.7%; Score 287; DB 4; Length 362;
Best Local Similarity 99.7%; Pred. NO. 1.7e-127;
Matches 337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGGAGTGGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 60
    |||||||
Db 18 TGGAGTGGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 77
    |||||||

QY 61 CTGAGGCAATGGGTGGAAGGAAGATGACTGCAGTCTCTGGAAGAAACAGACCACC 120
    |||||||
Db 78 CTGAGGCAATGGGTGGAAGGAAGATGACTGCAGTCTCTGGAAGAAACAGACCACC 137
    |||||||

QY 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAATT 180
    |||||||
Db 138 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAATT 197
    |||||||

QY 181 TTCTCGTGAAGCAAGACTGACTGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 240
    |||||||
Db 198 TTCTCGTGAAGCAAGACTGACTGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 257
    |||||||

QY 241 CCTGCTTCGGGACAGCAGCCCTGGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 300
    |||||||
Db 258 CCTGCTTCGGGACAGCAGCCCTGGGAGATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 317
    |||||||

QY 301 GAAACATATTGACCTCGGAGGACATCTATGAGAGCAC 338
    |||||||
Db 318 GAAACATATTGACCTCGGAGGACATCTATGAGAGCAC 355
    |||||||

RESULT 2
US-08-738-367-8
; Sequence 8, Application US/08738367
; Patent No. 5827688
; GENERAL INFORMATION:
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/738,367
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

;
; MOLECULE TYPE: CDNA
; US-08-738-367-8
Query Match      0.9%; Score 23; DB 1; Length 296;
Best Local Similarity 100.0%; Pred. NO. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2425 ATATTTTCTCTAAAAA 2447
    |||||||
Db 268 ATATTTTCTCTAAAAA 290
    |||||||

RESULT 3
US-09-156-425-1
; Sequence 1, Application US/09156425B
; Patent No. 5962671
; GENERAL INFORMATION:
; APPLICANT: Baker, Brenda F.
; APPLICANT: Cowser, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF FAN EXPRESSION
; FILE REFERENCE: RTS-0009
; CURRENT APPLICATION NUMBER: US/09/156,425B
; CURRENT FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 3380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(2766)
; US-09-156-425-1

Query Match      0.9%; Score 23; DB 2; Length 3380;
Best Local Similarity 100.0%; Pred. NO. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2425 ATATTTTCTCTAAAAA 2447
    |||||||
Db 3350 ATATTTTCTCTAAAAA 3372
    |||||||

RESULT 4
US-08-737-980-1
; Sequence 1, Application US/08737980
; Patent No. 5843773
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Apoptosis Regulating Gene
; NUMBER OF SEQUENCES: 2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,980
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: KR 1995-6266
; FILING DATE: 24-MAR-1995
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 755 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: gDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; DEVELOPMENTAL STAGE: Fetus at 22 weeks gestation
```

; TISSUE TYPE: Liver
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 35..559
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 713..718
; US-08-737-980-1

Query Match 0.9%; Score 22; DB 2; Length 755;
Best Local Similarity 100.0%; Pred. No. 0.68;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2426 TATTTTCTCTAAAAA 2447
Db 726 TATTTTCTCTAAAAA 747

RESULT 5
US-08-878-989-12
; Sequence 12, Application US/08878989
; Patent No. 5885803
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,989
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT06
; CLONE: 827431
; US-08-878-989-12

Query Match 0.9%; Score 21; DB 2; Length 1282;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 411 CTACACAGAGAGGATGCCAG 431
Db 521 CTACACAGAGAGGATGCCAG 541

RESULT 6
US-09-272-796-12
; Sequence 12, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT06
; CLONE: 827431
; US-09-272-796-12

Query Match 0.9%; Score 21; DB 4; Length 1282;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 411 CTACACAGAGAGGATGCCAG 431
Db 521 CTACACAGAGAGGATGCCAG 541

RESULT 7
US-08-149-103-2
; Sequence 2, Application US/08149103
; Patent No. 5750367

```
; GENERAL INFORMATION:
; APPLICANT: Lawrence C. B. Chan
; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN RECEPTORS
; TITLE OF INVENTION: AND METHODS FOR USE OF SUCH
; TITLE OF INVENTION: RECEPTORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LYON & LYON
; STREET: 611 West Sixth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,103
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 204/052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-149-103-2
;
; Query Match 0.8%; Score 20; DB 1; Length 3116;
; Best Local Similarity 100.0%; Pred. No. 6.2;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1309 GGGCCCTGTGGCTGCTC 1328
Db 39 GGGCCCTGTGGCTGCTC 58

RESULT 8
US-08-451-883-2
; Sequence 2, Application US/08451883
; Patent No. 5798209
; GENERAL INFORMATION:
; APPLICANT: Lawrence C. B. Chan
; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW DENSITY
; TITLE OF INVENTION: LIPOPROTEIN RECEPTORS AND METHODS FOR
; TITLE OF INVENTION: USE OF SUCH RECEPTORS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LYON & LYON
; STREET: 633 West Fifth Street, Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: IBM MS-DOS (Version 6.22)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,883
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/149,103
; FILING DATE: No. 5798209ember 8, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Knight, Matthew W.
; REGISTRATION NUMBER: 36,846
; REFERENCE/DOCKET NUMBER: 212/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-451-883-2
;
; Query Match 0.8%; Score 20; DB 1; Length 3116;
; Best Local Similarity 100.0%; Pred. No. 6.2;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1309 GGGCCCTGTGGCTGCTC 1328
Db 39 GGGCCCTGTGGCTGCTC 58

RESULT 9
US-08-456-647B-5/c
; Sequence 5, Application US/08456647B
; Patent No. 5811516
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,647B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/237,401
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07251/007002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
```

one

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3307 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
IMMEDIATE SOURCE:
CLONE: Tyro-3
FEATURE:
NAME/KEY: CDS
LOCATION: 237..2859
US-08-456-647B-5

Query Match 0.8%; Score 20; DB 1; Length 3307;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336
|||||
DB 2789 TGGCTGCTGCTCCAGCTGCC 2770

RESULT 10
US-08-237-401A-5/c
Sequence 5, Application US/08237401A
Patent No. 5837448
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/237,401A
FILING DATE: 02-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Halle Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07251/0007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3307 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
IMMEDIATE SOURCE:
CLONE: Tyro-3
FEATURE:
NAME/KEY: CDS
LOCATION: 237..2859
US-08-237-401A-5

Query Match 0.8%; Score 20; DB 2; Length 3307;
Best Local Similarity 100.0%; Pred. No. 6.2;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336
|||||
DB 2789 TGGCTGCTGCTCCAGCTGCC 2770

RESULT 11
US-08-445-640-9/c
Sequence 9, Application US/08445640
Patent No. 5709858
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J.
APPLICANT: Mark, Melanie R.
APPLICANT: Scadden, David T.
APPLICANT: Baker, Kevin P.
APPLICANT: Baron, Will F.
TITLE OF INVENTION: Protein Tyrosine Kinases
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,640
FILING DATE: 22-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/170558
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/157563
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 854C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3785 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-445-640-9

Query Match 0.8%; Score 20; DB 1; Length 3785;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336
|||||
DB 2632 TGGCTGCTGCTCCAGCTGCC 2613

RESULT 12
US-08-170-558-9/c
Sequence 9, Application US/08170558
Patent No. 6001621
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J.
APPLICANT: Mark, Melanie R.

```
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,558
; FILING DATE: 20-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C1
; TELEPHONE: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3785 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-170-558-9

Query Match 0.8%; Score 20; DB 3; Length 3785;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336
Db 2632 TGGCTGCTGCTCCAGCTGCC 2613

RESULT 13
US-08-447-314-9/c
; Sequence 9, Application US/08447314
; Patent No. 6087144
; GENERAL INFORMATION:
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,314
```

```
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C1D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/952-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3785 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-447-314-9

Query Match 0.8%; Score 20; DB 3; Length 3785;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336
Db 2632 TGGCTGCTGCTCCAGCTGCC 2613

RESULT 14
US-08-445-461-9/c
; Sequence 9, Application US/08445461
; Patent No. 6096527
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Scadden, David T.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Baron, Will F.
; TITLE OF INVENTION: Protein Tyrosine Kinases
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,461
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 854C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1896
```

Job time : 120 secs

; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3785 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-445-461-9

Query Match 0.8%; Score 20; DB 3; Length 3785;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1317 TGGCTGCTGCTCCAGCTGCC 1336
|||||
Db 2632 TGGCTGCTGCTCCAGCTGCC 2613

RESULT 15

US-08-466-120-1
; Sequence 1, Application US/08466120
; Patent No. 5869284
; GENERAL INFORMATION:
; APPLICANT: CAO, ET AL.
; TITLE OF INVENTION: Retinoic Acid Receptor Epsilon
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,120
; FILING DATE: June 6, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07266
; FILING DATE: 24 JUN 94
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-354
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1649 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; US-08-466-120-1

Query Match 0.8%; Score 19; DB 2; Length 1649;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2202 AGGGCTGCAAGGAATCTT 2220
|||||
Db 526 AGGGCTGCAAGGAATCTT 544

Search completed: March 15, 2003, 05:02:39

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:08:41 ; Search time 16 Seconds
(without alignments)
1371.239 Million cell updates/sec

Title: US-09-960-643-2
Perfect score: 2513
Sequence: 1 MGRKEEDDCSSWKQTNR.....VKASGSSHRAGQTGVCLIM 476

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2513	100.0	476	9	US-09-935-464-5
2	2362	94.0	460	9	US-09-935-464-3
3	1246	49.6	357	9	US-10-024-036B-2
4	1152.5	45.9	370	9	US-10-142-356-7
5	1152.5	45.9	370	10	US-09-817-181-4
6	1151	45.8	317	9	US-09-935-464-36
7	1016.5	40.4	326	10	US-09-817-181-2
8	944.5	37.6	280	10	US-09-835-788A-16
9	747.5	29.7	501	10	US-09-797-039-2
10	747.5	29.7	501	12	US-10-153-921-2
11	725.5	24.9	765	9	US-10-217-357-4
12	625.5	24.9	765	10	US-09-975-326-4
13	625.5	24.9	766	9	US-09-934-406-2
14	625.5	24.9	766	9	US-10-217-357-2
15	625.5	24.9	766	10	US-09-975-326-2
16	606.5	24.1	385	9	US-10-116-332-2
17	594.5	23.7	545	9	US-10-142-356-6
18	592	23.6	556	12	US-10-096-960-4
19	585.5	23.3	648	9	US-10-024-036B-5

20	585.5	23.3	817	10	US-09-992-481-4	Sequence 4, Appli
21	575.5	22.9	543	10	US-09-740-627-1	Sequence 1, Appli
22	571.5	22.7	565	12	US-10-096-960-2	Sequence 23, Appli
23	563.5	22.4	295	9	US-09-988-462-23	Sequence 16, Appli
24	558	22.2	821	9	US-10-081-119-16	Sequence 11, Appli
25	556.5	22.1	522	10	US-09-740-627-11	Sequence 25, Appli
26	542.5	21.6	463	9	US-09-988-462-25	Sequence 17, Appli
27	538.5	21.4	639	10	US-09-854-731-17	Sequence 5, Appli
28	531	21.1	514	10	US-09-922-138-5	Sequence 2, Appli
29	531	21.1	514	10	US-09-841-683-2	Sequence 5, Appli
30	530	21.1	454	10	US-09-771-161A-238	Sequence 238, App
31	514	20.5	298	10	US-09-858-664A-17	Sequence 17, Appli
32	513	20.4	740	10	US-09-771-161A-265	Sequence 265, App
33	513	20.4	740	10	US-09-771-161A-266	Sequence 266, App
34	508.5	20.2	265	9	US-09-764-868-814	Sequence 814, App
35	508.5	20.2	259	10	US-09-828-313-38	Sequence 38, Appli
36	506	20.1	549	10	US-09-828-313-39	Sequence 39, Appli
37	505.5	20.1	406	10	US-09-771-161A-210	Sequence 210, App
38	502.5	20.0	508	10	US-09-858-664A-18	Sequence 18, Appli
39	498	19.8	370	9	US-10-081-119-34	Sequence 34, Appli
40	496.5	19.8	625	10	US-09-854-731-18	Sequence 18, Appli
41	495.5	19.7	623	10	US-09-854-731-4	Sequence 4, Appli
42	490	19.5	464	9	US-09-988-462-22	Sequence 22, Appli
43	486	19.3	576	10	US-09-854-731-19	Sequence 19, Appli
44	483	19.2	387	10	US-09-771-161A-188	Sequence 188, App
45	480.5	19.1	459	10	US-09-740-627-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-09-935-464-5
; Sequence 5, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Harrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-464-5

Query Match	100.0%	Score 2513;	DB 9;	Length 476;
Best Local Similarity	100.0%;	Pred No. 1.3e-181;		
Matches 476;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy 1	MGRKEEDDCSSWKQTNRKKTIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKSPAF	60		
Db 1	MGRKEEDDCSSWKQTNRKKTIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKSPAF	60		
Qy 61	RDSLENEIAVLAKKHENIVLTLEDIYESTHYLYVMQVLSGGELFDRILRGVYTERDA	120		
Db 61	RDSLENEIAVLAKKHENIVLTLEDIYESTHYLYVMQVLSGGELFDRILRGVYTERDA	120		
Qy 121	SLVTIQVLSAVKYLHENGIVHVRDLKPENLIYLTPEPNSKIMITDFGLSKMEQNGIMSTAC	180		
Db 121	SLVTIQVLSAVKYLHENGIVHVRDLKPENLIYLTPEPNSKIMITDFGLSKMEQNGIMSTAC	180		
Qy 181	GTPGYVAPEVLAOKPYSKAVDCWSIGVITYILLCGYPPPYEETESKLFPEKIEGYEFES	240		
Db 181	GTPGYVAPEVLAOKPYSKAVDCWSIGVITYILLCGYPPPYEETESKLFPEKIEGYEFES	240		

Qy 241 PFWDISSAKDFCHLLEKDPNRYTCEKALSHPWIDGNTALHROIYPSVSLQIQNFA 300
Db 241 PFWDISSAKDFCHLLEKDPNRYTCEKALSHPWIDGNTALHROIYPSVSLQIQNFA 300
Qy 301 KSKWROAFNAVAHVHMRKLHNNLHSPGVRPEVNRPEPETOASETSPSSPEITTEAPV 360
Db 301 KSKWROAFNAVAHVHMRKLHNNLHSPGVRPEVNRPEPETOASETSPSSPEITTEAPV 360
Qy 361 LDHVSVALPALQPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHQSLAAGPCGCCS 420
Db 361 LDHVSVALPALQPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHQSLAAGPCGCCS 420
Qy 421 SCLNIGSKGKSSYCEPTLLKANKKONFKSEVMPVKASGSHCRAGOTGVCLIM 476
Db 421 SCLNIGSKGKSSYCEPTLLKANKKONFKSEVMPVKASGSHCRAGOTGVCLIM 476
RESULT 2
US-09-935-464-3
; Sequence 3, Application US/09935464
; Publication No. US20030027152A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-464-3

Query Match 94.0%; Score 2362; DB 9; Length 460;
Best Local Similarity 100.0%; Pred. No. 3e-170;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGRKEEDDCSSWKQTNIIRKTFIFMEVLGSGAFSEVFLVKQRLTKGLFALKCIKSPAF 60
Db 1 MGRKEEDDCSSWKQTNIIRKTFIFMEVLGSGAFSEVFLVKQRLTKGLFALKCIKSPAF 60
Qy 61 RDSSLENEIAVLKKIKHENIVTLEDIYESTHYLYMQLVSGGELFDRILERGVTTEKDA 120
Db 61 RDSSLENEIAVLKKIKHENIVTLEDIYESTHYLYMQLVSGGELFDRILERGVTTEKDA 120
Qy 121 SLVTOQVLSAVKYLHENGIVHRDLKPNLLYLTPPEENSKIMITDGLSKMEQNGIMSTAC 180
Db 121 SLVTOQVLSAVKYLHENGIVHRDLKPNLLYLTPPEENSKIMITDGLSKMEQNGIMSTAC 180
Qy 181 GTPGVVAPEVLAQPKYSKAVDCWSIGVITVILLCGYPFPFYEETESKLFKEIKEGYEFES 240
Db 181 GTPGVVAPEVLAQPKYSKAVDCWSIGVITVILLCGYPFPFYEETESKLFKEIKEGYEFES 240
Qy 241 PFWDISSAKDFCHLLEKDPNRYTCEKALSHPWIDGNTALHROIYPSVSLQIQNFA 300
Db 241 PFWDISSAKDFCHLLEKDPNRYTCEKALSHPWIDGNTALHROIYPSVSLQIQNFA 300
Qy 301 KSKWROAFNAVAHVHMRKLHNNLHSPGVRPEVNRPEPETOASETSPSSPEITTEAPV 360
Db 301 KSKWROAFNAVAHVHMRKLHNNLHSPGVRPEVNRPEPETOASETSPSSPEITTEAPV 360
Qy 361 LDHVSVALPALQPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHQSLAAGPCGCCS 420
Db 361 LDHVSVALPALQPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHQSLAAGPCGCCS 420

Qy 421 SCLNIGSKGKSSYCEPTLLKANKKQ 447
Db 421 SCLNIGSKGKSSYCEPTLLKANKKQ 447
RESULT 3
US-10-024-036B-2
; Sequence 2, Application US/10024036B
; Publication No. US20030028004A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 68730 and 69112, protein Kinase
; TITLE OF INVENTION: Molecules and Uses Therefor
; FILE REFERENCE: MPI2000-521P1R(W)
; CURRENT APPLICATION NUMBER: US/10/024,036B
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/258222
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-036B-2
Query Match 49.6%; Score 1246; DB 9; Length 357;
Best Local Similarity 73.1%; Pred. No. 2.6e-86;
Matches 237; Conservative 40; Mismatches 45; Indels 2; Gaps 2;
Qy 1 MGRKEEDDCSSWKQTNIIRKTFIFMEVLGSGAFSEVFLVKQRLTKGLFALKCI-KKSPA 59
Db 1 MARENGESSSSWKQOAEIDIKKIFETKLTGTGAFSEVLAEEKATGKLFVAVKCIKPKALK 60
Qy 60 FRDSSLENEIAVLKKIKHENIVTLEDIYESTHYLYMQLVSGGELFDRILERGVTTEKD 119
Db 61 GKSSLENEIAVLKKIKHENIVALEDIYESPNNHLYMQLVSGGELFDRILERGVTTEKD 120
Qy 120 ASLVTQVLSAVKYLHENGIVHRDLKPNLLYLTPPEENSKIMITDGLSKMEQNG-IMST 178
Db 121 ASTLIRQVLDVAYVLLHRRGIVHRDLKPNLLYYSQDEESKIMISDFGLSKMEKGDVYMT 180
Qy 179 ACGTPGVVAPEVLAQPKYSKAVDCWSIGVITVILLCGYPFPFYEETESKLFKEIKEGYEF 238
Db 181 ACGTPGVVAPEVLAQPKYSKAVDCWSIGVITVILLCGYPFPFYDENDSKLFEQILKAEYEF 240
Qy 239 ESPFWDDISSAKDFCHLLEKDPNRYTCEKALSHPWIDGNTALHROIYPSVSLQIQN 298
Db 241 DSPYWDIDSDSAKDFIRNLMEKDPNKRYTCEQAARHPWIAGDTALNKNIHESVSAQIRKN 300
Qy 299 FAKSKWROAFNAVAHVHMRKLHM 322
Db 301 FAKSKWROAFNATAVVHRMRKLHL 324
RESULT 4
US-10-142-356-7
; Sequence 7, Application US/10142356
; Publication No. US20030036183A1
; GENERAL INFORMATION:
; APPLICANT: Boylan, John F.
; APPLICANT: Bowers, Alex J.
; TITLE OF INVENTION: Serine-Threonine Kinase Member, h2520-40
; FILE REFERENCE: 01017/37177A
; CURRENT APPLICATION NUMBER: US/10/142,356
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 60/290,276
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 370
; TYPE: PRT


```

; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-788A-16

Query Match          37.6%; Score 944.5; DB 10; Length 280;
Best Local Similarity 77.5%; Pred. No. 9.7e-64;
Matches 176; Conservative 26; Mismatches 24; Indels 1; Gaps

Qy  97 MOLVSGGELFDRLRIGVYTEKASVLVQVLSAVKYLHENGIVHRDLKPENLLYLTPEE 156
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db  1  MOLVSGGELFDRIIVEKGFYTEKASTLIQIRVLDAVYVLRMGIVHRDLKPENLLYYSQDE 60
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy  157 NSKIMIDFGLSKMEQNG-TMSTACGTPGGVVAPEVLAAQKPYSKAVDCWSIGVITYILLCG 215
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  61 ESKIMISDFGLSKMEGKDVMSACGTPGGVVAPEVLAAQKPYSKAVDCWSIGVITYILLCG 120
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy  216 YPPPYEETSEKLPKIKEGYYEFESPWDDISSAKDFICHLEKDPNERYTCEKALSHP 275
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  121 YPPPYENDSKLFRQILKAETEFDSPYWDIDISAKDFIRNLMEKDPNKRYTCEQAARHP 180
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy  276 WIDGNTALHRDIYPSVLSIQKNFAKSKWRQAFNAAVVHHMKLHM 322
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db  181 WIAGDTALNKNIHESVSAQIRKNFAKSKWRQAFNATAVVRMRKRLHL 227
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 9
US-09-797-039-2
; Sequence 2, Application US/09797039
; Patent No. US20020042099A1
; GENERAL INFORMATION:
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 2504, 15977, AND 14760, NOVEL PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 10448-017001
; CURRENT APPLICATION NUMBER: US/09/797,039
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-797-039-2

Query Match          29.7%; Score 747.5; DB 10; Length 501;
Best Local Similarity 35.9%; Pred. No. 1.5e-48;
Matches 175; Conservative 81; Mismatches 176; Indels 55; Gaps

Qy  15 QTTNIRKTFIPMEVLGSGAFSEVFLVKORITGKLFALKCIKKSPAFRD-----SSLENEI 69
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db  16 QPSEVTDRIYDLGVIKITEEFCEIFRAKDKTGTGLKLTCKKFKQ-----RDGRKVRKAANEI 71
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy  70 AVLKIKHENIVTLIEDIYESTHYYLMVLVSGGELFDRILERCYVTEKASVLVQVLS 129
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  72 GILKMKVHPNTLQLVDFVTRKEYFIFLELATGREVDFWLDQGYYSERDTSNVVRQVLE 131
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy  130 AVKYLHENGIVHRDLKPENLLYLTPENSKITMTDFGLSKMEONGINSTAGPGYVAPE 189
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  132 AVAYLSUKIVHRNLKLENLYYVRLNKNKIVISDFHLAKLE-NGLIKFCGTPPEYLAPE 190
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy  190 VLAQKPYSKAVDCWSIGVITYILLCGYPPPYEETE-----SKLFEKIKEGYYEFSP 241
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  191 VVGRORYGRPVDCAIGVIMYILLSGNPPPYEEVEEDDYENHDKNLFRKILAGDYEFDSP 250
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Qy  242 FWDIDISAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVLSIQKNFAK 301
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```


Db 251 YWDDISQAADLVTRLMVEYEQDQRTAEBAISHEWISGNAASDKNIKDCVCAQIEKNFAR 310
QY 302 SKWRQAFNAAVVHHMRKILHMLNLSHSGVGRPEVENRPPETQASSETSPPSPEITITEAPVL 361
Db 311 AKWKKAQVRVTTL---MKRL-----RAPE--QSSTAAQASASATDTATPCA 350
QY 362 ---DHSVALPALQPCQGRRP-----TAPGRSLNCLVNGSL--HSSSLVPMHQSL 411
Db 351 AGGATAAAASGATSAPGDAARAASDNVAPADRSATPATDGSATPATDGSVTPATDGS 410
QY 412 AAGPCGCCSCLN-----IGSKGKSSYCSRPSTLLKKRANKKQNFSEVWV---PVKASGSSH 464
Db 411 TPATDGSVTPATDRSATPATDGRATPATEESTVPTTQSSAMLATKAAATPEPAMAQPDST 470
QY 465 CRAGQTG 471
Db 471 APEGATG 477

RESULT 10
US-10-153-921-2
; Sequence 2, Application US/10153921
; Patent No. US20020142430A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL000612DV
; CURRENT APPLICATION NUMBER: US/10/153,921
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/207,281
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 09/734,030
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: PRT
; ORGANISM: HOMO SAPIEN
US-10-153-921-2

Query Match 29.7%; Score 747.5; DB 12; Length 501;
Best Local Similarity 35.3%; Pred. No. 1.5e-48;
Matches 173; Conservative 86; Mismatches 170; Indels 61; Gaps 12;

QY 15 QTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKSPAPRD-----SSLENEI 69
Db 16 QPSEVTDRLDQGVIKTEEFCEIFRAKDKTTGKLTCKKFKQK----RDGRKVRKAAKNEI 71

QY 70 AVLKIKKHENIVTLEDIESTTHYVLVMOVLVSGGELFDRILRGVYVTEKDASLVITQOVL 129
Db 72 GILKMVKHPNQLQVDFVTRKEYFFILELATGREVFDWLDQGYYSERDTSNVVRQVLE 131

QY 130 AVLYLHENGIVHRDLKPENLLYLTPENSKIMITDPLGSKMONGTMSACGTGGVVAPE 189
Db 132 AVAYLHSLKLVHRNKLLENLVYTRNLRKNSKIVISDFHLAKLE--NLGIKEPCGTPEYLAPE 190

QY 190 VLAQKPYSAKVDSCWSIGVITYILLCGYPFPYEETE-----SKLFEKIEGYEFESF 241
Db 191 VVGRQRYGRPVDCAIGVIMYILSLGNPPPYEEVEEDDYENHDKNLFRKILAGDYEFDSP 250

QY 242 FWDDISAKDFCHLLEKDPNRYCEKALSHPWIDGNLTALHRDIPSVSLQIQKNFAK 301
Db 251 YWDDISQAADLVTRLMVEYEQDQRTAEBAISHEWISGNAASDKNIKDCVCAQIEKNFAR 310

QY 302 SKWRQAFNAAVVHHMRKILHMLNLSHSGVGRPEVENRPPETQAS--ETSRP-----SSPE 352
Db 311 AKWKKAQVRVTTLMKRLR-----APEQSTAAQASASATDTATPCAAGGATAAAA 359
QY 353 ITITEAPVLDHVALPALQPCQHGRRPTAPGGRSLNCLVNGSL--HSSSLVPMHQGS 410

Db 360 SCATSAPEGDAARAASDNVAPADRSATPATDG--SATPATDGSVTPATDGSITPATDGS 417
QY 411 L-----AAGPCGCCSCLNIGSKGKSSYCSRPSTLLKKRANKKQNFSEVWV---PVKASG 461
Db 418 VTPVTDRSATP-----ATDGRATPATEESTVPTTQSSAMLATKAAATPEPAMAQ 467
QY 462 SSHCRAGQTG 471
Db 468 DSTAPEGATG 477

RESULT 11
US-10-217-357-4
; Sequence 4, Application US/10217357
; Publication No. US20030023063A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; TITLE OF INVENTION: No. US20030023063A1el Human Kinases and Polynucleotides Encod
; FILE OF INVENTION: Same
; FILE REFERENCE: LEX-0254--USA
; CURRENT APPLICATION NUMBER: US/10/217,357
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US/09/975,326
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 60/239,821
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 765
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-217-357-4

Query Match 24.9%; Score 625.5; DB 9; Length 765;
Best Local Similarity 34.8%; Pred. No. 4.1e-39;
Matches 142; Conservative 82; Mismatches 143; Indels 41; Gaps 11;

QY 2 GRKEEDDCSSWK-----KOTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCI 54
Db 365 GGPELDRGISPEGVNGNRCSESTLLEKIKYKIGVDGNFVAVKRECIDRSTGKEFALKII 424

QY 55 KKSAPAF-RDSSLENEIAVLKKIKHENIVTLEDIESTTHYVLVMOVLVSGGELFDRILRG 113
Db 425 DKACCGGHELIENEVSILRRKHPNIIIMLVEEMETATELFLVWELVKGDLFPAITSST 484

QY 114 VYTEKASLVITQOVLASVAYKYLHENGIVHRDLKPENLLYLT-PEENSKIMITDFGLSKMEQ 172
Db 485 KYTERDGSAMVYNLANALRYLHGLSIVHRDIKPENLLVCEYPDGTTKSLKLGDFGLATVVE 544

QY 173 NGIMSTAGCTPGYVAPEVLAQKPYSAKVDCHSIVGIVITVILLCGYPFPYEET--ESKLPEK 230
Db 545 -GPLYTVCGTPTYVAPEIAETGYGLKYVDIWAAGVITVILLCGFPFPRSENQLQEDLFDQ 603

QY 231 IKEGYEFESFPDDDISAKDFICHLEKDPNRYCEKALSHPWIDGNLTALHRDIYPS 290
Db 604 ILAGKLEFPAPYWDNITOSAKELISQMLQVNEARCTAGQILSHPWVSDSDASQENMQAE 663

QY 291 VSLQIQKNF--AKSKWRQAFNAAVVHHMR-----KLHMLNLH-----SPGVRPEVENRPP 338
Db 664 VTGKLQKHFNALPKONSTTTTGVSVIMNTALDKEQIFCSKHODSGRPGMEP-ISPVP 722

QY 339 ETQASETSRPSPEITITEAPVLDHVALPALQPCQHGRRPTAPGG 386
Db 723 -----SVEEIPVGEAVPAPTTPESPSTPH-CPPAAPGG 754

RESULT 12
US-09-975-326-4
; Sequence 4, Application US/09975326
; Patent No. US20020128458A1
; GENERAL INFORMATION:

QY	2	GRKEDDCSSWK-----KQTNIRKTFIFMVELVSGAFSEVFLVKORLTGKLFALKCI	54
Db	366	GGPELDRCISPEGVNGNRCSESSTLLEKYYKIGKVIGDGNFAVVKCEIDSTGCKEFALKII	425
QY	55	KKSPAF-RDSSLENEIAVLKKIKHENIVTLEDIESTTHYYLMQVLSGGELFDRILERG	113
Db	426	DKAKCCCKEHLIENEVSILRRVKHPNIIIMLVEEMETATELFLVMEVLVKGDDLFDQ	485
QY	114	VYTEKDSLVITQOVL SAVKYLHENGIVHRDLKDPENLLYLT-PEENSKIMITDFGLSKMEQ	172
Db	486	KYTERDGSAMVYNLANALRYLHGLSIVHRDIKPENLLVCEYDGTGKSLKLGDFGLATVVE	545
QY	173	NGIMSTACGTGGVVAPEVLVLAQKPYSKAVDCWSGTGVTIYLLCGYPPFYET--ESKLF	230
Db	546	-GPLYTVCGTPTYVAPETIAETGYGLKVDIWAAGVITYILLCGFPFPRSENKLODQ	604
QY	231	IKEGYEFESFWDIDISESAKDFCHLLKDPNERVTCCKALSHPWIDGNTALHRDIYPS	290
Db	605	ILAKLEFPAPYDNIITDSKELLISQMLQVNVZARTAGQLISHPWSDASQENNMQAE	664
QY	291	VSLOIQKNF--AKSKWQAFNAAAVHHMR-----KLHMLNH-----SPGVRPEVENRPP	338
Db	665	VTGKLQHFNNALPKQNSTTGTGYSVIMNTALDKEGQIFCSKHQDSGRGPCMEP-ISPVP	723
QY	339	ETQASETSRSPSPITITEAPVLDSVALPALTLQPCQHGRRPTARGG	386
Db	724	-----SVEEIPVPGEAVPAPTPPESPSTPH-CPPAAPGG	755
RESULT 14			
US-10-217-357-2			
; Sequence 2, Application US/10217357			
; Publication No. US20030023063A1			
; GENERAL INFORMATION:			
; APPLICANT: Mathur, Brian			
; APPLICANT: Turner, C. Alexander Jr.			
; TITLE OF INVENTION: No. US20030023063A1el Human Kinases and Polynucleot			
; TITLE OF INVENTION: Same			
; FILE REFERENCE: LEX-0254-USA			
; CURRENT APPLICATION NUMBER: US/10/217,357			
; CURRENT FILING DATE: 2002-08-09			
; PRIOR APPLICATION NUMBER: US/09/975,326			
; PRIOR FILING DATE: 2001-10-11			
; PRIOR APPLICATION NUMBER: US 60/239,821			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 4			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 2			
; LENGTH: 766			
; TYPE: PRT			
; ORGANISM: homo sapiens			
US-10-217-357-2			
Query Match 24.9%; Score 625.5; DB 9; Length 766;			
Best Local Similarity 34.8%; Pred.No. 4.le-39;			
Matches 142; Conservative 82; Mismatches 143; Indels 41; Gaps			
QY	2	GRKEDDCSSWK-----KQTNIRKTFIFMVELVSGAFSEVFLVKORLTGKLFALKCI	54
Db	366	GGPELDRCISPEGVNGNRCSESSTLLEKYYKIGKVIGDGNFAVVKCEIDSTGCKEFALKII	425
QY	55	KKSPAF-RDSSLENEIAVLKKIKHENIVTLEDIESTTHYYLMQVLSGGELFDRILERG	113
Db	426	DKAKCCCKEHLIENEVSILRRVKHPNIIIMLVEEMETATELFLVMEVLVKGDDLFDQ	485
QY	114	VYTEKDSLVITQOVL SAVKYLHENGIVHRDLKDPENLLYLT-PEENSKIMITDFGLSKMEQ	172
Db	486	KYTERDGSAMVYNLANALRYLHGLSIVHRDIKPENLLVCEYDGTGKSKLGDGLATVVE	545
QY	173	NGIMSTACGTGGVVAPEVLVLAQKPYSKAVDCWSGTGVTIYLLCGYPPFYET--ESKLF	230
Db	546	-GPLYTVCGTPTYVAPETIAETGYGLKVDIWAAGVITYILLCGFPFPRSENKLODQ	604

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 14:17:20 ; Search time 566 Seconds
(without alignments)
9736.113 Million cell updates/sec

Title: US-09-960-643-1
Perfect score: 2447
Sequence: 1 tggagtgagctcaagcag.....ttttctctataaaaaaaa 2447

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq.101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2447	100.0	2447	24 AAD36140	Human calmodulin k
2	1687	68.9	1956	22 AAI60703	Human polynucleoti
3	1144	46.8	2165	22 AAI58917	Human polynucleoti
4	772	31.5	2689	22 AAS31014	Human diagnostic a
5	355	14.5	588	23 ABK43480	DNA encoding novel
6	287	11.7	362	20 AAX51838	Human secreted pro
7	274	11.2	476	22 AAS27154	cDNA encoding nove
8	274	11.2	476	23 ABK43797	DNA encoding novel
9	149	6.1	1361	22 AAH64901	Human secreted pro

10	107	4.4	373	24	ABL99884	Human secretory po
11	60	2.5	60	24	ABN35390	Human spliced tran
12	32	1.3	1512	23	AAS70157	DNA encoding novel
13	32	1.3	1774	23	AAS68767	DNA encoding novel
14	32	1.3	1774	23	AAS90919	DNA encoding novel
15	32	1.3	1774	23	AAS90919	DNA encoding novel
16	25	1.0	25	24	AAD36148	Human CAMK-X1 full
17	25	1.0	25	24	AAD36155	Human CAMK-X1 full
18	25	1.0	30	24	AAD36157	Human CAMK-X1 dele
19	25	1.0	30	24	AAD36158	Human CAMK-X1 dele
20	25	1.0	31	24	AAD36156	Human CAMK-X1 full
21	25	1.0	31	24	AAD36159	Human CAMK-X1 dele
22	23	0.9	85	22	AH36908	Human colon cancer
23	23	0.9	296	19	AAV04278	Secreted protein C
24	23	0.9	534	21	AAD00307	Arabidopsis SERK i
25	23	0.9	3380	20	AZ22338	DNA encoding a FAN
26	22	0.9	755	17	AAT41464	Human foetal liver
27	22	0.9	1420	22	AHL13806	Human cDNA sequenc
28	22	0.9	16759	24	ABL36297	Human lysosomal ac
29	21	0.9	1032	22	AAS06711	Polynucleotide seq
30	21	0.9	1282	20	AAK06835	Disease associated
31	21	0.9	18183	22	AAS46279	Tumour suppressor
32	21	0.9	18183	24	ABL70111	Chemically treated
33	20	0.9	18183	24	ABK31158	Signal transductio
34	20	0.8	352	22	AAI88729	Human polynucleoti
35	20	0.8	358	24	ABO85700	Arabidopsis thalia
36	20	0.8	445	22	AAK62638	Human immune/haema
37	20	0.8	445	22	AAK82954	Human immune/haema
38	20	0.8	445	22	AAK82955	Human immune/haema
39	20	0.8	1108	24	ABN98328	Arabidopsis thalia
40	20	0.8	1157	24	ABQ24014	Oligonucleotide fo
41	20	0.8	1157	24	ABQ24015	Oligonucleotide fo
42	20	0.8	2144	21	AAQ73212	Human 17723 recept
43	20	0.8	2334	22	AAQ74277	Rice calcium depen
44	20	0.8	2550	15	AAQ71333	Mature murine deve
45	20	0.8	3116	16	AAQ88688	Rat very low densi
	20	0.8	3307	19	AAV55888	Receptor protein t

ALIGNMENTS

RESULT 1
AAD36140
ID AAD36140 standard; DNA; 2447 BP.

AC AAD36140;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human calmodulin kinase, CAMK-X1 gene.

KW Human; cytostatic; antisense gene therapy; screening; protein kinase;
KW cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-X1;
KW calmodulin kinase; enzyme; gene; chromosome 1q32.1-32.3; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 70..1500
CDS
FT /*tag= a
FT /product= "Human CAMK-X1 protein"

PN WC200224947-A2.

XX 28-MAR-2002.

XX 20-SEP-2001; 2001WO-IB02237.

XX 20-SEP-2000; 2000US-233999P.

PR 02-OCT-2000; 2000US-237419P.

PR 02-OCT-2000; 2000US-237423P.

PR 04-OCT-2000; 2000US-238558P.

PR 10-MAY-2001; 2001US-290555P.

XX (KINE-) KINETEK PHARM INC.
PA (YFBR-) UNIV BRITISH COLUMBIA.
XX
PI Yoganathan T, Delaney AD;
DR WPI: 2002-394145/42.
DR P-PSDB: RAE22764.
XX
XX Diagnosing cancer, comprises determining the upregulation of expression
PT of a nucleic acid sequence encoding a protein kinase or upregulation of
PT expression of the protein kinase, in the cancer -
XX
PS Claim 16; Page 62-64; 87pp; English.
XX
CC The invention relates to a method for screening biologically active agent
CC that modulates cancer associated protein kinase function. The invention
CC also relates to a method for diagnosing cancer comprising determining the
CC upregulation of expression of a nucleic acid sequence encoding a protein
CC kinase. The method is useful for diagnosing cancer. A protein kinase is
CC useful for screening biological agents that modulate cancer associated
CC protein kinase function. Downregulating the activity of protein kinase is
CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon
CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy
CC derived tumours and inflammatory samples such as arthritic synovium, for
CC amplified DNA in the cell or increased expression of corresponding mRNA
CC or protein and is also useful to detect differences in expression levels
CC such as molecular weight, amino acid and nucleotide sequences between the
CC two cells. The present sequence is human calmodulin kinase CAMK-X1 gene
CC located on chromosome 1q32.1-32.3.
XX
SQ Sequence 2447 BP; 590 A; 707 C; 604 G; 546 T; 0 other;

Query Match 100.0%; Score 2447; DB 24; Length 2447;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGAGTGGGAGCTCAAGCAGGATTCCTCCCGAGTCCCTGGCATCCTCAGAAAGTTCACAT 60
DB 1 TGGAGTGGGAGCTCAAGCAGGATTCCTCCCGAGTCCCTGGCATCCTCAGAAAGTTCACAT 60
QY 61 CTGGAGGCAATGGGTGGAAGGAAGAAGATGACTGCTGAGTTCCTGGAGAAACAGACACAC 120
DB 61 CTGGAGGCAATGGGTGGAAGGAAGAAGATGACTGCTGAGTTCCTGGAGAAACAGACACAC 120
QY 121 AACATCCGGGAAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTCTCAGAAAGTT 180
DB 121 AACATCCGGGAAAACCTTCATTTTATGGAAGTGTGGGATCAGGAGCTTCTCAGAAAGTT 180
QY 181 TTCCTGTGGAAGCAAGCACTGACCTGGAAGTCTTTGCTCTGAAGTGCATCAGAAGTCA 240
DB 181 TTCCTGTGGAAGCAAGCACTGACCTGGAAGTCTTTGCTCTGAAGTGCATCAGAAGTCA 240
QY 241 CCTGCTCTCCGGGACAGCAGCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300
DB 241 CCTGCTCTCCGGGACAGCAGCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300
QY 301 GAAAACATTTGACCTTGGAGGACATCTATGAGAGCACCACCTACTACTCTGTCATG 360
DB 301 GAAAACATTTGACCTTGGAGGACATCTATGAGAGCACCACCTACTACTCTGTCATG 360
QY 361 CAGCTTGTCTGTGGGGAGCTTTTACCGGATCTGGAGCGGGTGTCTACACAGAG 420
DB 361 CAGCTTGTCTGTGGGGAGCTTTTACCGGATCTGGAGCGGGTGTCTACACAGAG 420
QY 421 AAGGATCCAGTCTGGTGATCCAGCAGGCTTGTTCGGCAGTGAATACCTACATGAGAT 480
DB 421 AAGGATCCAGTCTGGTGATCCAGCAGGCTTGTTCGGCAGTGAATACCTACATGAGAT 480
QY 481 GGCATCTCCACAGAGACTTAAAGCCGAAAACCTGCTTTACCTTACCCTGAGAGAAC 540
DB 481 GGCATCTCCACAGAGACTTAAAGCCGAAAACCTGCTTTACCTTACCCTGAGAGAAC 540

QY 541 TCTAAGATCATGATCAGTCACTTGGTCTGTCCAAAGATGGACAGAAATGCAATCATGTCC 600
DB 541 TCTAAGATCATGATCAGTCACTTGGTCTGTCCAAAGATGGACAGAAATGCAATCATGTCC 600
QY 601 ACTGCTCTGGGACCCAGGCTAGCTGCTCCAGAAAGTGTGGCCCAAGAAACCTACAGC 660
DB 601 ACTGCTCTGGGACCCAGGCTAGCTGCTCCAGAAAGTGTGGCCCAAGAAACCTACAGC 660
QY 661 AAGGCTGTGGATGCTGTGCTCATCAGCTACACCTACATATTTGCTGTGGATACCCC 720
DB 661 AAGGCTGTGGATGCTGTGCTCATCAGCTACACCTACATATTTGCTGTGGATACCCC 720
QY 721 CGGTTCTATGAAGAAACGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAG 780
DB 721 CGGTTCTATGAAGAAACGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAG 780
QY 781 TTTGAGTCTCCATTTCTGGATGACATTTCTGAGTCAGCCAAAGGACTTTATTTGCCACTTG 840
DB 781 TTTGAGTCTCCATTTCTGGATGACATTTCTGAGTCAGCCAAAGGACTTTATTTGCCACTTG 840
QY 841 CTTGAGAGGATCCGAAACGAGCTGACACCTGTGAGAGGCTTTGAGTCATCCTCGATT 900
DB 841 CTTGAGAGGATCCGAAACGAGCTGACACCTGTGAGAGGCTTTGAGTCATCCTCGATT 900
QY 901 GACGAAACACGGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAAG 960
DB 901 GACGAAACACGGCCCTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAAG 960
QY 961 RACITTTGATAGAGCAAGTGGAGGCAAGCCTTCAACCCAGCAGCTGTGGTGCACCATG 1020
DB 961 RACITTTGATAGAGCAAGTGGAGGCAAGCCTTCAACCCAGCAGCTGTGGTGCACCATG 1020
QY 1021 AGGAAGCTTACATGAACTGCACAGCCCGGGCTGCGCCAGAGTGGAGAAACAGGCCG 1080
DB 1021 AGGAAGCTTACATGAACTGCACAGCCCGGGCTGCGCCAGAGTGGAGAAACAGGCCG 1080
QY 1081 CCTGAACTCAAGCTCAGAAACCTTAGACCCAGCTCCCTGAGATCAGCTACCGGAG 1140
DB 1081 CCTGAACTCAAGCTCAGAAACCTTAGACCCAGCTCCCTGAGATCAGCTACCGGAG 1140
QY 1141 GCACCTGTCTGGACACAGTGTAGCAGTCCCTGCTGACCCCAATTTACCTGCGACAT 1200
DB 1141 GCACCTGTCTGGACACAGTGTAGCAGTCCCTGCTGACCCCAATTTACCTGCGACAT 1200
QY 1201 GCGCGCGGGCCACTGCCCCCTGGTGGAGGTCCTCAACTGCTGTGTCATAGCTCCCTC 1260
DB 1201 GCGCGCGGGCCACTGCCCCCTGGTGGAGGTCCTCAACTGCTGTGTCATAGCTCCCTC 1260
QY 1261 CACATCAGCAGCGCTGGTGGCCATGCATCAGGGGTCCCTGGCGCGCGGGCCCTGGGC 1320
DB 1261 CACATCAGCAGCGCTGGTGGCCATGCATCAGGGGTCCCTGGCGCGCGGGCCCTGGGC 1320
QY 1321 TGCTGCTCCAGTGCCTGAACTTTGGGACAAAGAAAGTCCCTACTGCTCTGAGCCC 1380
DB 1321 TGCTGCTCCAGTGCCTGAACTTTGGGACAAAGAAAGTCCCTACTGCTCTGAGCCC 1380
QY 1381 ACATCTCTAAAAGGGCAACAAAACAGAACTTCAAGTCGGAGGTCATGGTACCAAGTT 1440
DB 1381 ACATCTCTAAAAGGGCAACAAAACAGAACTTCAAGTCGGAGGTCATGGTACCAAGTT 1440
QY 1441 AAAGCAGTGGCAGCTCCCTGCGGCGAGGCGAGCTGGAGTCTGCTCATATTATGTA 1500
DB 1441 AAAGCAGTGGCAGCTCCCTGCGGCGAGGCGAGCTGGAGTCTGCTCATATTATGTA 1500
QY 1501 TTCTGGAGCTGTGCTGCTATGCTCAATTTTCAGGAGACATATTCAACTCCTCTGCT 1560
DB 1501 TTCTGGAGCTGTGCTGCTATGCTCAATTTTCAGGAGACATATTCAACTCCTCTGCT 1560
QY 1561 CTTCCAAACCTGTGCTATCCGCGAGGAGGAGGAGGAGGAGCAAGTGGAGGAGGCTT 1620
DB 1561 CTTCCAAACCTGTGCTATCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGCTT 1620
QY 1621 AGCAGGAGCAGTCTTCTGGCCAGAAAGCAGCTGTGCTGCCAGCGGGGCGCCCTCATAG 1680

QY 312 GACCTGGAGGACATCTATGAGAGCACCACCACTACTACCTGGTTCATGCAGCTGTTTC 371
|||||
Db 307 GACCTGGAGGACATCTATGAGAGCACCACCACTACTACCTGGTTCATGCAGCTGTTTC 366
|||||
QY 372 TGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGTGTCTACAGAGAAGATGCCAG 431
|||||
Db 367 TGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGTGTCTACAGAGAAGATGCCAG 426
|||||
QY 432 TCTGGTGATCCAGCAGGCTTGTGGCAGTGAATACCTACATGATGATGCGATCGTCA 491
|||||
Db 427 TCTGGTGATCCAGCAGGCTTGTGGCAGTGAATACCTACATGATGATGCGATCGTCA 486
|||||
QY 492 CAGAGACTTAAAGCCGGAACCTGCTTTACCTTACCCTGAAGAGAAGCTCTAAGATCAT 551
|||||
Db 487 CAGAGACTTAAAGCCGGAACCTGCTTTACCTTACCCTGAAGAGAAGCTCTAAGATCAT 546
|||||
QY 552 GATCACTGACTTTGGTCTCTCAAGATGGAACAGAAATGGCATATGTCCATGCCGTGG 611
|||||
Db 547 GATCACTGACTTTGGTCTCTCAAGATGGAACAGAAATGGCATATGTCCATGCCGTGG 606
|||||
QY 612 GACCCAGGCTACGTGGCTCCAGAGTGTGCTGGCCAGAAACCTACAGCAAGGCTGTGA 671
|||||
Db 607 GACCCAGGCTACGTGGCTCCAGAGTGTGCTGGCCAGAAACCTACAGCAAGGCTGTGA 666
|||||
QY 672 TTGCTGGTCCATCGGCGTCATCACCTACATATTGCTGTGGATACCCCGCTTCTATGA 731
|||||
Db 667 TTGCTGGTCCATCGGCGTCATCACCTACATATTGCTGTGGATACCCCGCTTCTATGA 726
|||||
QY 732 AGAAGCGGAGTCTAAGCTTTTCAGAGATCAAGAGGGCTACTATGATTTGAGTTCCTCC 791
|||||
Db 727 AGAAGCGGAGTCTAAGCTTTTCAGAGATCAAGAGGGCTACTATGATTTGAGTTCCTCC 786
|||||
QY 792 ATTCTGGGATGACATTTCTGAGTCAAGCAAGGACATTTATTGGCCACTTGCCTGGAAGA 851
|||||
Db 787 ATTCTGGGATGACATTTCTGAGTCAAGCAAGGACATTTATTGGCCACTTGCCTGGAAGA 846
|||||
QY 852 TCGAAGCGAGGCTACACCTGTGAGAAGGCTTTGAGTCATCCCTGGATTGACGGAACAC 911
|||||
Db 847 TCGAAGCGAGGCTACACCTGTGAGAAGGCTTTGAGTCATCCCTGGATTGACGGAACAC 906
|||||
QY 912 GGCCCTCCACGGGACATCTACCATCAGTCAGCTCAGATCCAGATCCAGAGAAGCTTTGCTAA 971
|||||
Db 907 AGCCCTCCACGGGAGATCTACCCATCAGTCAGCTCAGCTCCAGATPCAGAGAAGCTTTGCTAA 966
|||||
QY 972 GAGCAAGTGGAGCAAGCTTCAACGACGACGCTGTGTGTGACCAACATCAGGAGGCTACA 1031
|||||
Db 967 GAGCAAGTGGAGCAAGCTTCAACGACGACGCTGTGTGTGACCAACATCAGGAGGCTACA 1026
|||||
QY 1032 CATGAACCTGACAGCCCGGGCGTCGCCCCAGAGGTGGAGAACAGGCCCTGAAACTCA 1091
|||||
Db 1027 CATGAACCTGACAGCCCGGGCGTCGCCCCAGAGGTGGAGAACAGGCCCTGAAACTCA 1086
|||||
QY 1092 AGCCTCAGAAACCTCTAGACCCAGCTCCCTCAGATCAACATCAGGAGGACCTGCTCT 1151
|||||
Db 1087 AGCCTCAGAAACCTCTAGACCCAGCTCCCTCAGATCAACATCAGGAGGACCTGCTCT 1146
|||||
QY 1152 GGACCACTGTAGCACTCCCTGCCCTGACCCAAATTAACCTGCCAGCTGGCCGCGGCC 1211
|||||
Db 1147 GGACCACTGTAGCACTCCCTGCCCTGACCCAAATTAACCTGCCAGCTGGCCGCGGCC 1206
|||||
QY 1212 CACTGCCCTGTGGCAGGTCCCTCACTGCTCAATGGCTCCCTCCACATCAGCAG 1271
|||||
Db 1207 CACTGCCCTGTGGCAGGTCCCTCACTGCTCAATGGCTCCCTCCACATCAGCAG 1266
|||||
QY 1272 CAGCCTGGTGCCCATCATCAGGGTCCCTGGCCCGGGGCTGTGGCTGCTGCTCCAG 1331
|||||
Db 1267 CAGCCTGGTGCCCATCATCAGGGTCCCTGGCCCGGGGCTGTGGCTGCTGCTCCAG 1326
|||||
QY 1332 CTGCCCTGAACATTTGGGAGCAAGAAAGTCTCTACTGCTCTGAGCCACACTCCTCAA 1391
|||||
Db 1327 CTGCCCTGAACATTTGGGAGCAAGAAAGTCTCTACTGCTCTGAGCCACACTCCTCAA 1386
|||||

QY 1392 AAAGGCCAACAAAAACAGAACTTCAAGTCGAGGTCATGGTACCAGTTAAAGCCAGTGG 1451
|||||
Db 1387 AAAGGCCAACAAAAACAGAACTTCAAGTCGAGGTCATGGTACCAGTTAAAGCCAGTGG 1446
|||||
QY 1452 CAGTCCCACTGCGGGCAGGAGACTGGAGTCTGTCTCATTTATGTGATTCCTGGAGCC 1511
|||||
Db 1447 CAGTCCCACTGCGGGCAGGAGACTGGAGTCTGTCTCATTTATGTGATTCCTGGAGCC 1506
|||||
QY 1512 TGTGCTCATGTCACTGCAATTTTCAGGAGACATATCAACTCTCTGCTCTTCCAAACCT 1571
|||||
Db 1507 TGTGCTCATGTCACTGCAATTTTCAGGAGACATATCAACTCTCTGCTCTTCCAAACCT 1566
|||||
QY 1572 GGTGCTCATCCGCGAGAGGAGGAGCAGAGCAAGTGGAGCAGGGCTTAGCAGGAGCAG 1631
|||||
Db 1567 GGTGCTCATCCGCGAGAGGAGGAGCAGAGCAAGTGGAGCAGGGCTTAGCAGGAGCAG 1626
|||||
QY 1632 TTTCTGCCAGAGACACAGCCTGTGCCAGCGGGGAGGCCCTCATAGGAGGCCAGGA 1691
|||||
Db 1627 TTTCTGCCAGAGACACAGCCTGTGCCAGCGGGGAGGCCCTCATAGGAGGCCAGGA 1686
|||||
QY 1692 GGGAGCCCCAAGCGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACCA 1751
|||||
Db 1687 GGGAGCCCCAAGCGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACCA 1746
|||||
QY 1752 GCTTCCAGGTCTCCCTGACCTGCTCTATGCCCCACACCCCTACGTG 1800
|||||
Db 1747 GCTTCCAGGTCTCCCTGACCTGCTCTATGCCCCACACCCCTACGTG 1795
|||||

RESULT 3

AA158917
ID AA158917 standard; cdna; 2165 BP.
XX AC AA158917;
XX AC AA158917;
DF 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 1120.
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX Homo sapiens.
OS Homo sapiens.
XX WO200153312-A1.
PN WO200153312-A1.
XX 26-JUL-2001.
PD 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
PF 26-DEC-2000; 2000US-0488725.
XX 21-JAN-2000; 2000US-0552317.
PR 23-APR-2000; 2000US-0598042.
PR 09-JUL-2000; 2000US-0620312.
PR 19-JUL-2000; 2000US-0653450.
PR 03-AUG-2000; 2000US-0662191.
PR 14-SEP-2000; 2000US-0693036.
PR 19-OCT-2000; 2000US-0727344.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB; AAM39761.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT

PT such as central nervous system injuries -
PS Claim 1; SEQ ID NO 1120; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA38642-AA42213), with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 2165 BP; 533 A; 613 C; 563 G; 455 T; 1 other;
Query Match 46.8%; Score 1144; DB 22; Length 2165;
Best Local Similarity 99.8%; Pred No. 0;
Matches 1294; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 504 GCCGAAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTAAGATCATGACTGACTT 563
DB 708 GCCGAAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTAAGATCATGACTGACTT 767
QY 564 TGCTCTGTCACAGATGACACAAATGGCATCATGTCCTGCTGGAGCCCGAGCTA 623
DB 768 TGGTCTGTCCAGATGGAACAAATGGCATCATGTCCTGCTGGAGCCCGAGGATA 827
QY 624 CGTGGCTCCAGAAAGTCTGGCCAGAAACCTACAGAAAGCTGTGGATTGCTGGTCCAT 683
DB 828 CGTGGCTCCAGAAAGTCTGGCCAGAAACCTACAGAAAGCTGTGGATTGCTGGTCCAT 887
QY 684 CGCGTCATCACCATACATATGCTCTGTGATACCCCGTCTCTATGAAGAACGGAGTC 743
DB 888 CGGCGTCATCACCATACATATGCTCTGTGATACCCCGTCTCTATGAAGAACGGAGTC 947
QY 744 TAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGTTTGAGTCTCCATTCTGGGATGA 803
DB 948 TAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGTTTGAGTCTCCATTCTGGGATGA 1007
QY 804 CATTTCTGAGTCAGCCAAAGGACTTTATTTGCCACTTGCTTGAGAAGGATCCGAACGAGCG 863
DB 1008 CATTTCTGAGTCAGCCAAAGGACTTTATTTGCCACTTGCTTGAGAAGGATCCGAACGAGCG 1067
QY 864 GTACACCTGTGAGAAGGCTTGAGTCATCCCTGGATTGACGGAAACACGGCCCTCCACCG 923
DB 1068 GTACACCTGTGAGAAGGCTTGAGTCATCCCTGGATTGACGGAAACACAGCCCTCCACCG 1127
QY 924 GGACATCTACCCATCAGTCAGTCAGCTCCAGATCCAGAAAGACTTTGCTTAAGACCAAGTGAG 983
DB 1128 GGACATCTACCCATCAGTCAGCTCCAGATCCAGAAAGACTTTGCTTAAGACCAAGTGAG 1187
QY 984 GCAAGCCTTCAACGACAGCAGCTGTGTCACCATGAGGAAGCTACACATGAACCTGCA 1043
DB 1188 GCAAGCCTTCAACGACAGCAGCTGTGTCACCATGAGGAAGCTACACATGAACCTGCA 1247
QY 1044 CAGCCCGGGGTCGCGCCAGAGGTGGAGAACAGGCGCTGAACTCAAGCCTCAAGAAC 1103
DB 1248 CAGCCCGGGGTCGCGCCAGAGGTGGAGAACAGGCGCTGAACTCAAGCCTCAAGAAC 1307
QY 1104 CTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACCTGCTCGGACCAAGTGT 1163
DB 1308 CTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACCTGCTCGGACCAAGTGT 1367
QY 1164 AGCACTCCCTGCGCTGACCAAAATACCTGTCACAGATGCCCGCGGCCCACTGCCCTGG 1223

DB 1368 AGCACTCCCTGCGCTGACCAAAATTACCTGCGCAGCATGGCGCGGCCCACTGCCCTCGG 1427
QY 1224 TGGCAGGTCCTCACTCACTGCTGCTCAATGGCTCCCTCCACATCAGCAGCAGCTGGTGGC 1283
DB 1428 TGGCAGGTCCTCACTCACTGCTGCTCAATGGCTCCCTCCACATCAGCAGCAGCTGGTGGC 1487
QY 1284 CATGATCAGGCGTCCCTGCGCGCGCGGCTGTGGCTGTGCTGCCAGCTCCCTGAACAT 1343
DB 1488 CATGATCAGGCGTCCCTGCGCGCGCGGCTGTGGCTGTGCTGCCAGCTCCCTGAACAT 1547
QY 1344 TGGGACCAAGAAAGTCTCTACTGCTCTGAGCCCACTCTCTCAAAAAGGCCAACAA 1403
DB 1548 TGGGACCAAGAAAGTCTCTACTGCTCTGAGCCCACTCTCTCAAAAAGGCCAACAA 1607
QY 1404 AAACAGAACTTCAAGTCGGAGGTATGATACCACTTAAAGCCAGTGGCAGCTCCCACTG 1463
DB 1608 AAACAGAACTTCAAGTCGGAGGTATGATACCACTTAAAGCCAGTGGCAGCTCCCACTG 1667
QY 1464 CCGGCGAGGCGAGCTGGAGTCTGTCTATGATGATTCCTGAGGCTGTGCCATATGTC 1523
DB 1668 CCGGCGAGGCGAGCTGGAGTCTGTCTATGATGATTCCTGAGGCTGTGCCATATGTC 1727
QY 1524 ACTGCAATTTTCAGGAGACATATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCCG 1583
DB 1728 ACTGCAATTTTCAGGAGACATATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCCG 1787
QY 1584 GCAGGGAGGAGGAGGAGCAAGTGGAGCGGCTTAGCAGGAGCAGTTTCTGGCCAGA 1643
DB 1788 GCAGGGAGGAGGAGGAGCAAGTGGAGCGGCTTAGCAGGAGCAGTTTCTGGCCAGA 1847
QY 1644 AGCACCAGCCTGCTGCGAGCGGCGAGCCCTCATAGGAGGCCAGAGGAGGCCCAAG 1703
DB 1848 AGCACCAGCCTGCTGCGAGCGGCGAGCCCTCATAGGAGGCCAGAGGAGGCCCAAG 1907
QY 1704 GCGTAGAGCCTTGTGTAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGTGT 1763
DB 1908 GCGTAGAGCCTTGTGTAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGTGT 1967
QY 1764 CCTGACCTGCGCTGCTATGCCACACACCTTACGTC 1800
DB 1968 CCTGACCTGCGCTGCTATGCCACACACCTTACGTC 2004
RESULT 4
AAS31014
ID AAS31014 standard; cdNA; 2689 BP.
XX
AC AAS31014;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human diagnostic and therapeutic polynucleotide (DITHP) #29.
XX
KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;
KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;
KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;
KW respiratory disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200162927-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US06059.
XX
PR 24-FEB-2000; 2000US-0184693.
PR 24-FEB-2000; 2000US-0184697.
PR 24-FEB-2000; 2000US-0184698.
PR 24-FEB-2000; 2000US-0184768.
PR 24-FEB-2000; 2000US-0184769.
PR 24-FEB-2000; 2000US-0184770.
PR 24-FEB-2000; 2000US-0184771.

PR 24-FEB-2000; 2000US-0184772.
PR 24-FEB-2000; 2000US-0184773.
PR 24-FEB-2000; 2000US-0184774.
PR 24-FEB-2000; 2000US-0184776.
PR 24-FEB-2000; 2000US-0184777.
PR 24-FEB-2000; 2000US-0184797.
PR 24-FEB-2000; 2000US-0184813.
PR 24-FEB-2000; 2000US-0184837.
PR 24-FEB-2000; 2000US-0184841.
PR 24-FEB-2000; 2000US-0185213.
PR 24-FEB-2000; 2000US-0185216.
PR 12-MAY-2000; 2000US-0203785.
PR 15-MAY-2000; 2000US-0204226.
PR 16-MAY-2000; 2000US-0204525.
PR 16-MAY-2000; 2000US-0204821.
PR 16-MAY-2000; 2000US-0204908.
PR 16-MAY-2000; 2000US-0205232.
PR 17-MAY-2000; 2000US-0204815.
PR 17-MAY-2000; 2000US-0204863.
PR 17-MAY-2000; 2000US-0205221.
PR 17-MAY-2000; 2000US-0205285.
PR 17-MAY-2000; 2000US-0205286.
PR 17-MAY-2000; 2000US-0205287.
PR 17-MAY-2000; 2000US-0205323.
PR 17-MAY-2000; 2000US-0205324.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
PI Chen A, D'Sa SA, Anshay S, Dahl CR, Dam TC, Daniels SE;
PI Dufour GE, Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL;
PI Liu TF, Roseberry AM, Russo BH, Stockdreher TK, Daffo A;
PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;
PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;
XX
XX WPI: 2001-502867/55.
DR P-PSDB; AAU19443.
XX
XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.
PT enzymes, hormones and receptors, useful in diagnostics and therapeutics
PT -
XX
XX Claim 1: Page 310-311; 522pp; English.

CC The invention relates to polynucleotides (I) encoding diagnostic and
CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes,
CC and proteins involved in growth and development and receptors. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate DITHP expression. For example, (I) and
CC (II) may be used to treat disorders associated with decreased polypeptide
CC expression by rectifying mutations or deletions in a patient's genome,
CC that affect the activity of the DITHPs, by expressing inactive proteins
CC or supplementing the patient's own production of them. (I) and (II)
CC may be used to treat diseases, for example, cell proliferative disorder,
CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,
CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,
CC (I) may be used to produce the DITHPs, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the protein. (I) and
CC its complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative
CC therapy. (II) may also be used as antigens in the production of
CC antibodies against DITHPs and in assays to identify modulators of DITHP
CC expression and activity. The anti-DITHP antibodies and antagonists may
CC also be used to down regulate expression and activity. The anti-DITHP
CC antibodies may also be used as diagnostic agents for detecting the
CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbant
CC assay (ELISA)). AAS30986-AAS31196 represent human diagnostic and
CC therapeutic (DITHP) polynucleotides of the invention.
XX
XX Sequence 2689 BP; 642 A; 748 C; 678 G; 621 T; 0 other;

Query Match 31.5%; Score 772; DB 22; Length 2689;

		Best Local Similarity 99.9%; Pred. No. 0;				Matches 822; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	39	GGCATCTCTCAGAAAGCTTCAACTCTGGAGGCAATGGGTGCGAAAGAGAGATGACTGCAG	98						
Db	206	GGCATCTCTCAGAAAGCTTCAACTCTGGAGGCAATGGGTGCGAAAGAGAGATGACTGCAG	265						
Qy	99	TTCTGTGGAAGAAACAGACACCACACATCCGGAACCTTCATTTTATGAACTGCTGGG	158						
Db	266	TTCTGTGGAAGAAACAGACACCACACATCCGGAACCTTCATTTTATGAACTGCTGGG	335						
Qy	159	ATCAGGAGCTTTCTCAGAAAGTTTTCCTGGTGAAGCAAGACTGACTGGGAAGCTCTTTGC	218						
Db	326	ATCAGGAGCTTTCTCAGAAAGTTTTCCTGGTGAAGCAAGACTGACTGGGAAGCTCTTTGC	385						
Qy	219	TCTGAAGTGCATCAAGAAGTCACTGCTTCCGGGACAGCAGCTGGAGAAATGAGATTGC	278						
Db	386	TCTGAAGTGCATCAAGAAGTCACTGCTTCCGGGACAGCAGCTGGAGAAATGAGATTGC	445						
Qy	279	TGTCTTGAAAAAGATCAAGCATGAAACATTTGACCTGGAGACATCTATGAGAGCAC	338						
Db	446	TGTCTTGAAAAAGATCAAGCATGAAACATTTGACCTGGAGACATCTATGAGAGCAC	505						
Qy	339	CACCCACTACTACCTGCTCATGAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCT	398						
Db	506	CACCCACTACTACCTGCTCATGAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCT	565						
Qy	399	GGACGGGTGCTTACACAGAGAGGATGCCAGTCTGTGTCATCCAGCAGCTCTTGTCGGC	458						
Db	566	GGACGGGTGCTTACACAGAGAGGATGCCAGTCTGTGTCATCCAGCAGCTCTTGTCGGC	625						
Qy	459	AGTGAATATCTACATGAGAATGGCATCGTCCACAGAGACTTAAAGCCCGAAACCTGCT	518						
Db	626	AGTGAATATCTACATGAGAATGGCATCGTCCACAGAGACTTAAAGCCCGAAACCTGCT	685						
Qy	519	TTACTTTACCCCTCAAGAGAACTCTAAGATCATGATCAGTCTGCTGTGTCCTGTCAGAGT	578						
Db	686	TTACTTTACCCCTCAAGAGAACTCTAAGATCATGATCAGTCTGCTGTGTCCTGTCAGAGT	745						
Qy	579	GGAACAGAAATGGCATATGTCCTGCTGTGGACCCAGGCTAGCTGGCTCCAGAAAT	638						
Db	746	GGAACAGAAATGGCATATGTCCTGCTGTGGACCCAGGCTAGCTGGCTCCAGAAAT	805						
Qy	639	GCTGGCCCAAGAAACCTTACAGCAAGGCTGTGATGCTGCTCCATCGGCTCATACCTA	698						
Db	806	GCTGGCCCAAGAAACCTTACAGCAAGGCTGTGATGCTGCTCCATCGGCTCATACCTA	865						
Qy	699	CATATTGCTCTGTGTATACCCCGTCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAA	758						
Db	866	CATATTGCTCTGTGTATACCCCGTCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAA	925						
Qy	759	GATCAAGAGGGCTTACATGATGTTGAGTCTCCATCTCTGGGATGACATTTCTGAGTCAGC	818						
Db	926	GATCAAGAGGGCTTACATGATGTTGAGTCTCCATCTCTGGGATGACATTTCTGAGTCAGC	985						
Qy	819	CAAGGACTTTATTTGCCACTTGCTTGAAGAGGATCCGAACGAG	861						
Db	986	CAAGGACTTTATTTGCCACTTGCTTGAAGAGGATCCGAACGAG	1028						

RESULT 5
ABK43480
ID ABK43480 standard; cDNA; 588 BP.

XX AC ABK43480;

XX DT 05-JUN-2002 (first entry)

XX DE DNA encoding novel central nervous system protein #60.

XX KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;

KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200155318-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01332.
XX
31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-581633/65.
 DR P-PSDB; AAU87150.
 XX
 XX New isolated nucleic acid encoding a protein for diagnosing,
 PT preventing, treating or ameliorating medical conditions and used as
 PT food additives or preservatives -
 XX
 XX Claim 1; SEQ ID NO 70; 837pp; English.
 XX
 XX The invention describes an isolated nucleic acid molecule (I) encoding a
 CC novel central nervous system protein. (I) and polypeptides (III) encoded
 CC by (I), are used to treat a medical conditions and in diagnosis of a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
 CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
 CC adenocarcinomas and irritable bowel syndrome, reproductive system
 CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
 CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
 CC leukaemia, disorders involving neovascularisation e.g. malignancies,
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
 CC acute kidney failure and blood related disorders e.g. myocardial
 CC infarction. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,
 Query Match 14.5%; Score 355; DB 23; Length 588;
 Best Local Similarity 100.0%; Pred. No. 2.5e-162;
 Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TGGAGTGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 60
 Db 145 TGGAGTGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 204
 Qy 61 CTGGAGGCAATGGCTGAAAGGAGAGAGTACTGCTAGTTCTCTGGAGAAACAGACACC 120
 Db 205 CTGGAGGCAATGGCTGAAAGGAGAGAGTACTGCTAGTTCTCTGGAGAAACAGACACC 264
 Qy 121 AACATCCGGAACCTTCATTTTATGGAAGTGGTGGGATCAGAGCTTCTCAGAAGTT 180
 Db 265 AACATCCGGAACCTTCATTTTATGGAAGTGGTGGGATCAGAGCTTCTCAGAAGTT 324
 Qy 181 TTCTGGTGAAGCAAGAGTACTGCTGGAAGTCTTTCTGCTGAAGTGCATCAAGAAGTCA 240
 Db 325 TTCTGGTGAAGCAAGAGTACTGCTGGAAGTCTTTCTGCTGAAGTGCATCAAGAAGTCA 384
 Qy 241 CCTGCCCTCCGGGACAGACGCTGGAGAGATGAGATTGCTGTGTTGAAAGAGATCAAGCAT 300
 Db 385 CCTGCCCTCCGGGACAGACGCTGGAGAGATGAGATTGCTGTGTTGAAAGAGATCAAGCAT 444
 Qy 301 GAAACATTTGACCTTGGAGGACATCTATGAGAGCACCACCACCTACTACTGG 355
 Db 445 GAAACATTTGACCTTGGAGGACATCTATGAGAGCACCACCACCACCTACTACTGG 499

RESULT 6
 AAX51838
 ID AAX51838 standard; DNA; 362 BP.
 XX
 AC AAX51838;
 XX
 DT 22-JUN-1999 (first entry)
 XX
 DE Human secreted protein 5' EST SEQ ID NO: 52.
 XX
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO9906552-A2.
 XX
 PD 11-FEB-1999.
 XX
 PF 31-JUL-1998; 98WO-IB01236.
 XX
 PR 01-AUG-1997; 97US-0905223.
 XX
 PA (GEST) GENSET.
 XX
 PI Duclert A, Dumas Milne Edwards J, Lacroix B;
 XX
 DR WPI; 1999-153782/13.
 DR P-PSDB; AAY13038.
 XX
 XX New isolated brain-derived nucleic acids - used to develop products
 PT which may have cytokine, immune, regulatory, haematopoiesis
 PT regulating, anti-inflammatory or tumour inhibition activity
 XX
 PS Claim 1; Page 212; 577pp; English.
 XX
 XX AAX51787 to AAX52019 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12987 to
 CC AAY13219, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX
 SQ Sequence 362 BP; 107 A; 79 C; 99 G; 76 T; 1 other;
 Query Match 11.7%; Score 287; DB 20; Length 362;
 Best Local Similarity 99.7%; Pred. No. 3.2e-129;
 Matches 337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 TGGAGTGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 60
 Db 18 TGGAGTGGAGCTCAAGCAGGATTTCTCCGAGTCCCTGGCATCTCAGAGCTTCAACT 77
 Qy 61 CTGGAGGCAATGGCTGAAAGGAGAGTACTGCTAGTTCTCTGGAGAAACAGACACC 120
 Db 78 CTGGAGGCAATGGCTGAAAGGAGAGTACTGCTAGTTCTCTGGAGAAACAGACACC 137
 Qy 121 AACATCCGGAACCTTCATTTTATGGAAGTGGTGGGATCAGAGCTTCTCAGAAGTT 180

```
|||||
Db 138 AACATCGGAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 197
QY 181 TTCCTGCTGAAGCAAGACTGACTGGAGCTCTTTGCTCTGAAGTGCATCAAGAACTCA 240
Db 198 TTCCTGCTGAAGCAAGACTGACTGGAGCTCTTTGCTCTGAAGTGCATCAAGAACTCA 257
QY 241 CCTGCCCTCCGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300
Db 258 CCTGCCCTCCGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCAT 317
QY 301 GAAAACATTGTGACCTGGAGACATCTATGAGAGCAC 338
Db 318 GAAAACATTGTGACCTGGAGACATCTATGAGAGCAC 355

RESULT 7
AAS27154
ID AAS27154 standard; cDNA; 476 bp.
AC AAS27154;
XX
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 189.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01312.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0228868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246534.
PR 08-NOV-2000; 2000US-0246535.
PR 08-NOV-2000; 2000US-0246536.
PR 08-NOV-2000; 2000US-0246537.
PR 08-NOV-2000; 2000US-0246538.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246540.
PR 08-NOV-2000; 2000US-0246541.
PR 08-NOV-2000; 2000US-0246542.
PR 08-NOV-2000; 2000US-0246543.
PR 08-NOV-2000; 2000US-0246544.
PR 08-NOV-2000; 2000US-0246545.
PR 08-NOV-2000; 2000US-0246546.
PR 08-NOV-2000; 2000US-0246547.
PR 08-NOV-2000; 2000US-0246548.
PR 08-NOV-2000; 2000US-0246549.
PR 08-NOV-2000; 2000US-0246550.
PR 08-NOV-2000; 2000US-0246551.
PR 08-NOV-2000; 2000US-0246552.
PR 08-NOV-2000; 2000US-0246553.
PR 08-NOV-2000; 2000US-0246554.
PR 08-NOV-2000; 2000US-0246555.
PR 08-NOV-2000; 2000US-0246556.
PR 08-NOV-2000; 2000US-0246557.
PR 08-NOV-2000; 2000US-0246558.
PR 08-NOV-2000; 2000US-0246559.
PR 08-NOV-2000; 2000US-0246560.
PR 08-NOV-2000; 2000US-0246561.
PR 08-NOV-2000; 2000US-0246562.
PR 08-NOV-2000; 2000US-0246563.
PR 08-NOV-2000; 2000US-0246564.
PR 08-NOV-2000; 2000US-0246565.
PR 08-NOV-2000; 2000US-0246566.
PR 08-NOV-2000; 2000US-0246567.
PR 08-NOV-2000; 2000US-0246568.
PR 08-NOV-2000; 2000US-0246569.
PR 08-NOV-2000; 2000US-0246570.
PR 08-NOV-2000; 2000US-0246571.
PR 08-NOV-2000; 2000US-0246572.
PR 08-NOV-2000; 2000US-0246573.
PR 08-NOV-2000; 2000US-0246574.
PR 08-NOV-2000; 2000US-0246575.
PR 08-NOV-2000; 2000US-0246576.
PR 08-NOV-2000; 2000US-0246577.
PR 08-NOV-2000; 2000US-0246578.
PR 08-NOV-2000; 2000US-0246579.
PR 08-NOV-2000; 2000US-0246580.
PR 08-NOV-2000; 2000US-0246581.
PR 08-NOV-2000; 2000US-0246582.
PR 08-NOV-2000; 2000US-0246583.
PR 08-NOV-2000; 2000US-0246584.
PR 08-NOV-2000; 2000US-0246585.
PR 08-NOV-2000; 2000US-0246586.
PR 08-NOV-2000; 2000US-0246587.
PR 08-NOV-2000; 2000US-0246588.
PR 08-NOV-2000; 2000US-0246589.
PR 08-NOV-2000; 2000US-0246590.
PR 08-NOV-2000; 2000US-0246591.
PR 08-NOV-2000; 2000US-0246592.
PR 08-NOV-2000; 2000US-0246593.
PR 08-NOV-2000; 2000US-0246594.
PR 08-NOV-2000; 2000US-0246595.
PR 08-NOV-2000; 2000US-0246596.
PR 08-NOV-2000; 2000US-0246597.
PR 08-NOV-2000; 2000US-0246598.
PR 08-NOV-2000; 2000US-0246599.
PR 08-NOV-2000; 2000US-0246600.
```


PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225216.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225271.
PR 14-AUG-2000; 2000US-0225275.
PR 14-AUG-2000; 2000US-0225278.
PR 14-AUG-2000; 2000US-0225279.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
FA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX P-PSDB; AAU87467.
XX WPI; 2001-581633/65.
XX P-PSDB; AAU87467.
XX
XX New isolated nucleic acid encoding a protein for diagnosing, and used as
XX preventing, treating or ameliorating medical conditions and used as
XX food additives or preservatives -
XX
XX Claim 1; SEQ ID No 387; 837pp; English.
XX
XX The invention describes an isolated nucleic acid molecule (I) encoding a
XX novel central nervous system protein. (I) and polypeptides (III) encoded
XX by (I), are used to treat a medical conditions and in diagnosis of a
XX pathological condition. Disorders which are diagnosed or treated include
XX autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
XX disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
XX angiogenesis, nervous system disorders e.g. Alzheimer's disease and
XX amyotrophic lateral sclerosis, infections caused by bacteria, viruses
XX e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
XX e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
XX adenocarcinomas and irritable bowel syndrome, reproductive system
XX disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
XX and pituitary dwarfism, cancers and disorders at the cellular level e.g.
XX leukaemia, disorders involving neovascularisation e.g. malignancies,

CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 11.2%; Score 274; DB 23; Length 476;
Best Local Similarity 100.0%; Pred. No. 6.9e-133;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGAGTGGGAGCTCAAGCAGGATTCTCCGAGTCCCTGGCATCTCAGAACTTCAACT 60
Db 145 TGGAGTGGGAGCTCAAGCAGGATTCTCCGAGTCCCTGGCATCTCAGAACTTCAACT 204

Qy 61 CTGGAGGCAATGGTGCAGGAAGATGATGACTGCAGTCTCTGGAAGAAACAGACCACC 120
Db 205 CTGGAGGCAATGGTGCAGGAAGATGATGACTGCAGTCTCTGGAAGAAACAGACCACC 264

Qy 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAACTT 180
Db 265 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAACTT 324

Qy 181 TTCTCTGGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 240
Db 325 TTCTCTGGTGAAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 384

Qy 241 CTGCGCTCTCCGGACAGCAGCCTGGGAGATGAGA 274
Db 385 CTGCGCTCTCCGGACAGCAGCCTGGGAGATGAGA 418

RESULT 9
AAH64901
ID AAH64901 standard; cdNA; 1361 BP.
AC AAH64901;
XX
XX
DT 11-SEP-2001 (first entry)
XX
XX Human secreted protein cDNA, SEQ ID NO: 177.
DE
DE Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200142451-A2.
XX
XX 14-JUN-2001.
XX
XX 07-DEC-2000; 2000WO-IB01938.
XX
XX 08-DEC-1999; 99US-0169629.
PR 06-MAR-2000; 2000US-0187470.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Bouqueleret L, Jobert S;
PI WPI; 2001-367870/38.
DR P-PSDB; AAG89298.
XX
XX Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
XX
XX Claim 7; Page 741-742; 921pp; English.
PS
XX The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides

CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET nucleic acid of the invention.
XX

SO Sequence 1361 BP; 438 A; 245 C; 302 G; 375 T; 1 other;

Query Match 6.1%; Score 149; DB 22; Length 1361;
Best Local Similarity 100.0%; Pred. No. 5.3e-62;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2299 TTCCTAATCGCTCCTGCTGCTCCCTTTCTCTCTGAAAGTCCACACACCATCTTGTCC 2358
Db 1177 TTCCTAATCGCTCCTGCTGCTCCCTTTCTCTCTGAAAGTCCACACACCATCTTGTCC 1236

Qy 2359 TTCCTAATCGCTCCTGCTCCTCCACCTCCAGCTTCATGCTGCTGCTTAAATAA 2418
Db 1237 TTCCTAATCGCTCCTGCTCCTCCACCTCCAGCTTCATGCTGCTGCTTAAATAA 1296

Qy 2419 ATGGACATATTTTCTCTTAAATAA 2447
Db 1297 ATGGACATATTTTCTCTTAAATAA 1325

RESULT 10
ABL99884
ID ABL99884 standard; cdNA; 373 BP.
AC ABL99884;
XX
XX 03-OCT-2002 (first entry)
DT
DE Human secretory polynucleotide (sptm) 139.
XX
XX Human; ss; gene; secretory protein; secretory polynucleotides; SPTM;
KW SPTM-related disease; somatic gene therapy; germline gene therapy;
KW severe combined immunodeficiency; intracellular parasite protection;
KW fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
KW immune disorder; AIDS; neurological disorder; Parkinson's disease;
KW motor neuron disorder; demyelinating disease; multiple sclerosis;
KW meningitis; abscess; prion diseases; cerebral palsy;
KW neuroskeletal disorder; peripheral nervous system disorder;
KW dermatomyositis; polymyositis; myopathy; myasthenia gravis;
KW mental disorder; Tourette's syndrome.
XX
XX Homo sapiens.
OS
XX WO200220756-A2.
PN
XX 14-MAR-2002.
PD
XX 30-AUG-2001; 2001WO-US27297.
XX
XX 05-SEP-2000; 2000US-229747P.
PR 05-SEP-2000; 2000US-229748P.
PR 05-SEP-2000; 2000US-229749P.
PR 05-SEP-2000; 2000US-229750P.
PR 05-SEP-2000; 2000US-229751P.
PR 05-SEP-2000; 2000US-230016P.
PR 05-SEP-2000; 2000US-230583P.
PR 06-SEP-2000; 2000US-230505P.


```
PR 06-SEP-2000; 2000US-230514P.
PR 06-SEP-2000; 2000US-230515P.
PR 06-SEP-2000; 2000US-230517P.
PR 06-SEP-2000; 2000US-230518P.
PR 06-SEP-2000; 2000US-230519P.
PR 06-SEP-2000; 2000US-230519P.
PR 06-SEP-2000; 2000US-230595P.
PR 06-SEP-2000; 2000US-230596P.
PR 06-SEP-2000; 2000US-230597P.
PR 06-SEP-2000; 2000US-230599P.
PR 06-SEP-2000; 2000US-230610P.
PR 06-SEP-2000; 2000US-230864P.
PR 06-SEP-2000; 2000US-230865P.
PR 06-SEP-2000; 2000US-230888P.
PR 06-SEP-2000; 2000US-230989P.
PR 06-SEP-2000; 2000US-230990P.
PR 06-SEP-2000; 2000US-230996P.
PR 07-SEP-2000; 2000US-230897P.
PR 07-SEP-2000; 2000US-230951P.
PR 07-SEP-2000; 2000US-231163P.
PR 07-SEP-2000; 2000US-231832P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
XX Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
PI Momiya MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
PI Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
XX
XX WPI: 2002-315658/35.
XX P-PSDB; ABB97887.
XX
XX Polynucleotide sequences encoding human secretory proteins useful for
PT gene therapy of e.g. genetic deficiency disorders, cancers, and
PT diseases caused by intracellular parasites -
XX
XX Claim 1; Page 331; 585pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
CC secretory (SPM) proteins. The SPM DNA and amino acid sequences are
CC useful for treating a disease or condition associated with the expression
CC of functional SPM. The SPM DNA sequences are useful for somatic or
CC germline gene therapy to correct a genetic deficiency (e.g. severe
CC combined immunodeficiency). The SPM DNA sequences are also useful in
CC providing protection against intracellular parasites (e.g. fungal
CC parasites and protozoan parasites). The SPM DNA and protein sequences
CC are also useful for diagnosing cell proliferative disorders, cancer,
CC immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
CC disease), motor neuron disorders, demyelinating diseases (e.g. multiple
CC sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,
CC neuroskeletal disorders, peripheral nervous system disorders,
CC dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental
CC disorders (e.g. Tourette's syndrome). cDNA sequences ABL99746 - ABL99929
CC represent human secretory polynucleotides of the invention.
XX
XX Sequence 373 BP; 92 A; 84 C; 98 G; 99 T; 0 other;
SQ
Query Match 4.4%; Score 107; DB 24; Length 373;
Best Local Similarity 99.4%; Pred. No. 1.5e-41;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 39 GGCATCCTCAGAAAGCTTCAACTCGGAGCAATGGGTGGAAGGAAGAGATGACTGCAG 98
DB 205 GGCATCCTCAGAAAGCTTCAACTCGGAGCAATGGGTGGAAGGAAGAGATGACTGCAG 264
QY 99 TTCTGTGGAAGAACAGACACCAACATCCGGAAAACTTCATTTTATGGAAGTCTGGG 158
DB 265 TTCTGTGGAAGAACAGACACCAACATCCGGAAAACTTCATTTTATGGAAGTCTGGG 324
QY 159 ATCAGGAGCTTTCAGAAAGTTTTCCTGGTGAAGCAA 196
DB 325 ATCAGGAGCTTTCAGAAAGTTTTCCTGGTGAAGCAA 362
```

RESULT 11

```
ABN35390
ID ABN35390 standard; DNA; 60 BP.
XX
AC ABN35390;
XX
XX 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:8138.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
XX WC200210449-A2.
XX
XX 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-1B01903.
XX
XX 28-JUL-2000; 2000US-221607P.
XX 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI
XX WPI: 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
XX Example 1; SEQ ID 8138; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcripts. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 60 BP; 13 A; 18 C; 10 G; 19 T; 0 other;
SQ
Query Match 2.5%; Score 60; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1526 TGCAATTTTCAGGAGACATATTCAACTCCCTGCTCTTCCAAACCTGGTGTCTATCGGC 1585
DB 1 TGCAATTTTCAGGAGACATATTCAACTCCCTGCTCTTCCAAACCTGGTGTCTATCGGC 60
```

```
RESULT 12
AAS70157
ID AAS70157 standard; cDNA; 1512 BP.
XX
AC AAS70157;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #5961.
XX
Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
XX
DR P-PSDB; ABG05970.
XX
New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 5961; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1512 BP; 343 A; 391 C; 470 G; 308 T; 0 other;

Query Match 1.3%; Score 32; DB 23; Length 1512;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 683
|
DB 598 CCCTACAGCAAGGCTGTGGATTGCTGGTCCAT 629

RESULT 13
```

```
AAS68767
ID AAS68767 standard; cDNA; 1774 BP.
XX
AC AAS68767;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4571.
XX
Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
XX
DR P-PSDB; ABG04580.
XX
New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 4571; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1774 BP; 421 A; 488 C; 477 G; 388 T; 0 other;

Query Match 1.3%; Score 32; DB 23; Length 1774;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 CCTACAGCAAGGCTGTGGATTGCTGGTCCAT 683
|
DB 995 CCCTACAGCAAGGCTGTGGATTGCTGGTCCAT 1026

RESULT 14
AAS90919
ID AAS90919 standard; cDNA; 1774 BP.
```

```
XX AC AAS90919;
XX DT 13-FEB-2002 (first entry)
XX DE
XX KW DNA encoding novel human diagnostic protein #26723.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG26732.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity
XX PS Claim 1; SEQ ID No 26723; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1774 BP; 421 A; 488 C; 477 G; 388 T; 0 other;

Query Match 1.3%; Score 32; DB 23; Length 1774;
Best Local Similarity 100.0%; Pred. No. 5e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 CCCTACAGCAAGGCTGGGATTGCTGGTCCAT 683
Db ||||||||||||||||||||||||||||
995 CCCTACAGCAAGGCTGGGATTGCTGGTCCAT 1026

RESULT 15
AAD36148/C
ID AAD36148 standard; DNA; 25 BP.
XX AC AAD36148;
```

```
XX DT 09-AUG-2002 (first entry)
XX DE Human CAMK-X1 full length cDNA constructing 5' RACE PCR primer.
XX KW Human; cytostatic; antisense gene therapy; screening; protein kinase;
XX KW cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-X1;
XX KW calmodulin kinase; enzyme; RACE; rapid amplification of cDNA end; primer;
XX KW PCR; ss.
XX OS Homo sapiens.
XX PN WO200224947-A2.
XX PD 28-MAR-2002.
XX PF 20-SEP-2001; 2001WO-IB02237.
XX PR 20-SEP-2000; 2000US-233999p.
XX PR 02-OCT-2000; 2000US-237419p.
XX PR 02-OCT-2000; 2000US-237423p.
XX PR 04-OCT-2000; 2000US-238558p.
XX PR 10-MAY-2001; 2001US-290555p.
XX PA (KINE-) KINETEK PHARM INC.
XX PA (UYER-) UNIV BRITISH COLUMBIA.
XX PI Yoganathan T, Delaney AD;
XX WPI; 2002-394145/42.
XX PT Diagnosing cancer, comprises determining the upregulation of expression
XX PT of a nucleic acid sequence encoding a protein kinase or upregulation of
XX PT expression of the protein kinase, in the cancer
XX PS Example 2; Page 82; 87pp; English.
XX CC The invention relates to a method for screening biologically active agent
XX CC that modulates cancer associated protein kinase function. The invention
XX CC also relates to a method for diagnosing cancer comprising determining the
XX CC upregulation of expression of a nucleic acid sequence encoding a protein
XX CC kinase. The method is useful for diagnosing cancer. A protein kinase is
XX CC useful for screening biological agents that modulate cancer associated
XX CC protein kinase function. Downregulating the activity of protein kinase is
XX CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon
XX CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy
XX CC derived tumours and inflammatory samples such as arthritic synovium, for
XX CC amplified DNA in the cell or increased expression of corresponding mRNA
XX CC or protein and is also useful to detect differences in expression levels
XX CC such as molecular weight, amino acid and nucleotide sequences between the
XX CC two cells. The present sequence is 5' rapid amplification of cDNA ends
XX CC (RACE) PCR primer used for constructing human calmodulin kinase CAMK-X1
XX CC full length cDNA.
XX CC Note: The present sequence referred as SEQ ID NO: 17 is also shown in
XX CC page 25 of the specification but lacks two bases at the 5' end.
XX SQ Sequence 25 BP; 7 A; 2 C; 14 G; 2 T; 0 other;

Query Match 1.0%; Score 25; DB 24; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2358 CTTCCCCAGTTTCTCGCCCTCCAC 2382
Db ||||||||||||||||||||||||
25 CTTCCCCAGTTTCTCGCCCTCCAC 1

Search completed: March 14, 2003, 17:35:56
Job time : 573 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 15:22:10 ; Search time 7161 Seconds
(without alignments)
9944.779 Million cell updates/sec

Title: US-09-960-643-1
Perfect score: 2447
Sequence: 1 tggagtgggagctcaagcag.....tttttcttaaaaaaaaaa 2447

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	2447	100.0	2447	6	AX399682	AX399682 Sequence
2	2396	97.9	2474	9	BC032787	BC032787 Homo sapi
3	2143	87.6	2612	9	AK095713	AK095713 Homo sapi
4	2048	83.7	2464	9	AF428261	AF428261 Homo sapi
5	1636	66.9	1738	9	HS272L161	AL049688 Human gen
6	903	36.9	157875	9	HS272L16	AL023754 Human DNA
C 7	78	3.2	1249	9	HUMDIS70R	L26254 Human dinuc
8	71	2.9	2416	10	BC021840	BC021840 Mus muscu
9	71	2.9	2427	10	AF428262	AF428262 Mus muscu
10	62	2.5	182054	10	AL365314	AL365314 Mouse DNA
C 11	62	2.5	184754	2	AC022675	AC022675 Mus muscu
12	59	2.4	1013	10	D86557	D86557 Rattus norv
13	40	1.6	178272	2	AC126166	AC126166 Rattus no
C 14	40	1.6	191584	2	AC120477	AC120477 Rattus no
15	32	1.3	1448	10	BC014825	BC014825 Mus muscu
16	32	1.3	1480	9	HUMCK1	L41816 Homo sapien
C 17	32	1.3	2689	9	AK094026	AK094026 Homo sapi
C 18	32	1.3	155313	2	AC068315	AC068315 Homo sapi
19	32	1.3	157007	2	AC066599	AC066599 Homo sapi
20	32	1.3	161903	2	AC021996	AC021996 Homo sapi
21	32	1.3	185061	9	AC022382	AC022382 Homo sapi
C 22	32	1.3	189430	2	AC011610	AC011610 Homo sapi
C 23	32	1.3	197360	2	AC018829	AC018829 Homo sapi
C 24	27	1.1	191584	2	AC120477	AC120477 Rattus no
C 25	25	1.0	25	6	AX399696	AX399696 Sequence
C 26	25	1.0	25	6	AX399703	AX399703 Sequence
27	25	1.0	30	6	AX399708	AX399708 Sequence
28	25	1.0	30	6	AX399709	AX399709 Sequence
C 29	25	1.0	31	6	AX399704	AX399704 Sequence
C 30	24	1.0	30	6	AX399710	AX399710 Sequence
31	23	0.9	296	6	AR050391	AR050391 Sequence
32	23	0.9	534	6	AX024570	AX024570 Sequence
33	23	0.9	3380	6	AR078528	AR078528 Sequence
34	23	0.9	3380	9	HSFAN	X96586 H.sapiens m
C 35	23	0.9	146872	2	AC128229	AC128229 Rattus no
C 36	23	0.9	167898	2	AC103328	AC103328 Rattus no
37	22	0.9	755	6	AR062879	AR062879 Sequence
38	22	0.9	824	9	HSGRSPROT	Y09397 H.sapiens m
39	22	0.9	829	9	BC016281	BC016281 Homo sapi
C 40	22	0.9	860	9	HS325898	AJ325898 Homo sapi
C 41	22	0.9	1420	9	AK001010	AK001010 Homo sapi
C 42	22	0.9	2160	10	MM021392	U21392 Mus musculu
C 43	22	0.9	2347	9	AK055324	AK055324 Homo sapi
44	22	0.9	2462	9	AK097626	AK097626 Homo sapi
45	22	0.9	2648	10	BC031180	BC031180 Mus muscu

ALIGNMENTS

RESULT 1
AX399682
LOCUS AX399682
DEFINITION Sequence 3 from Patent WO0224947.
ACCESSION AX399682
VERSION AX399682.1
KEYWORDS GI:21335455
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Delaney,A.D. and Voganathan,T.
Cancer associated protein kinases and their uses
Patent: WO 0224947-A 3 28-MAR-2002;

AX399682 2447 bp DNA linear PAT 06-JUN-2002
Sequence 3 from Patent WO0224947.
AX399682.1 GI:21335455

FEATURES	KINETEK PHARMACEUTICALS INC (CA); UNIV BRITISH COLUMBIA (CA)									
	source	Location/Qualifiers 1..2447 /organism="Homo sapiens" /db_xref="taxon:9606" 70..1500 /note="unnamed protein product" /codon_start=1 /protein_id="CAD33521.1" /db_xref="GI:21335456"								
CDS	/translation="MKRKEEDDCSSWKKTNNIRKTFIPFMEVLGSGAFSEVFLVKORL TGLKFLKICGTGKSPAFRDSLENIKIHENIVTLEDIYESTHYLYLVMLVSG GELFDRLRGVVTEDASLVIQOVLSSAVKYLHENGIVHRDLKPENLLYLTPEENSKI MITDFGLSMEQINGINSTAGCTPGYVAPVLAQPKYSKAVDCWSIGIVYIILICGVPP FYETRSKLFPEKIEGYEPESPWDDISESAKDFICHLEKDPNERVTCERALSHPW IDGNTALHRDIYFSVSIQKNFAKSKWRQAFNAAVVHHMKRLHNLNLSPPGVREYE NRPEPQASSTSPSPSEITTEAPVLDHSVALPALTQLPCQHRPTAPGGRSLNCL VNGSLHISSLVPMHOGSLAAGPCGCCSLNLTGSKGSSYCSSEPTLLKANKKNQFK SEVMVPVKASGSHCRAGOTGVCLIM"									
BASE COUNT	590 a	707 c	604 g	546 t						
ORIGIN										
	Query Match	100.0%;	Score 2447;	DB 6;	Length 2447;					
	Best Local Similarity	100.0%;	Pred. No. 0;							
	Matches 2447;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
Qy	1	TGGAGTGGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT	60							
Db	1	TGGAGTGGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT	60							
Qy	61	CTGGAGGCAATGGGTGCGAAAGAAAGATGACTGCGAGTTCCTTGGAAAGAACAGACCACC	120							
Db	61	CTGGAGGCAATGGGTGCGAAAGAAAGATGACTGCGAGTTCCTTGGAAAGAACAGACCACC	120							
Qy	121	AACATCCGAAACCTTCATTTTATGGAAGTGTGGGATCAGAGGCTTCTCAGAAGTT	180							
Db	121	AACATCCGAAACCTTCATTTTATGGAAGTGTGGGATCAGAGGCTTCTCAGAAGTT	180							
Qy	181	TTCTGTGGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAAGTGCATCAAGAAGTCA	240							
Db	181	TTCTGTGGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAAGTGCATCAAGAAGTCA	240							
Qy	241	CCTGCCTTCGCGGACAGAGCCTTGAGAAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT	300							
Db	241	CCTGCCTTCGCGGACAGAGCCTTGAGAAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT	300							
Qy	301	GAACAATTTGACCTCGGAGACATCTATGAGAGCACCCACTACTACTCTGGTCTATG	360							
Db	301	GAACAATTTGACCTCGGAGACATCTATGAGAGCACCCACTACTACTCTGGTCTATG	360							
Qy	361	CAGCTTGTTCCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAG	420							
Db	361	CAGCTTGTTCCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAG	420							
Qy	421	AAGGATGCCAGTCTGGTGATCCAGAGCTTGTGTCGGAGTGAATACCTACATGAGAAT	480							
Db	421	AAGGATGCCAGTCTGGTGATCCAGAGCTTGTGTCGGAGTGAATACCTACATGAGAAT	480							
Qy	481	GGCATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC	540							
Db	481	GGCATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC	540							
Qy	541	TCTAAGATCATGATGACTTGGTGTGTGTCGAAGATGGAAACAGAAATGGCATGTCC	600							
Db	541	TCTAAGATCATGATGACTTGGTGTGTGTCGAAGATGGAAACAGAAATGGCATGTCC	600							
Qy	601	ACTGCCTGTGGACCCAGGCTACGTGGCTCCAGAAGTGTGGCCCGAGAACCCCTACAGC	660							
Db	601	ACTGCCTGTGGACCCAGGCTACGTGGCTCCAGAAGTGTGGCCCGAGAACCCCTACAGC	660							
Qy	661	AAGGCTGTGGATTGCTGGTCCATCGGCGTCAATCACTACATATTGCTGTGGGATACCCC	720							
Db	661	AAGGCTGTGGATTGCTGGTCCATCGGCGTCAATCACTACATATTGCTGTGGGATACCCC	720							

Qy	721	CCGTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTTACTATGAG	780	
Db	721	CCGTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTTACTATGAG	780	
Qy	781	TTTGAAGTCTCCATTCTGGGATGACATTTCTGAGTCAGGCCAAGGACTTTATTTGGCCACTTG	840	
Db	781	TTTGAAGTCTCCATTCTGGGATGACATTTCTGAGTCAGGCCAAGGACTTTATTTGGCCACTTG	840	
Qy	841	CTTGAGAAGGATTCGGAACAGCGGTACACCTCTGAGAAGGCCCTTGAGTCATCCCTGGATT	900	
Db	841	CTTGAGAAGGATTCGGAACAGCGGTACACCTCTGAGAAGGCCCTTGAGTCATCCCTGGATT	900	
Qy	901	GACGGAACACAGCCCTCCACCGGGACATCTACCCATCAGTCAGCCTCCAGATCCAGAG	960	
Db	901	GACGGAACACAGCCCTCCACCGGGACATCTACCCATCAGTCAGCCTCCAGATCCAGAG	960	
Qy	961	AACTTTGCTTAAGAGCAAGTGGAGGCAAGCTTCAACGCGACAGCTGTGGTGACCCATG	1020	
Db	961	AACTTTGCTTAAGAGCAAGTGGAGGCAAGCTTCAACGCGACAGCTGTGGTGACCCATG	1020	
Qy	1021	AGGAAGCTACACATGAACCTGCACAGCCCGGGCTCCGCGGAGAGTGGAGAACAGGCCG	1080	
Db	1021	AGGAAGCTACACATGAACCTGCACAGCCCGGGCTCCGCGGAGAGTGGAGAACAGGCCG	1080	
Qy	1081	CCTGAAACTCAAGCCTTCAGAAACCTCTAGACCCAGCTCCCTCGAGATCAACCATCCAG	1140	
Db	1081	CCTGAAACTCAAGCCTTCAGAAACCTCTAGACCCAGCTCCCTCGAGATCAACCATCCAG	1140	
Qy	1141	GCACCTGTCTGGACACAGTGTAGACACTCCCTGCCCTGACCCAAATACCTTGCACGAT	1200	
Db	1141	GCACCTGTCTGGACACAGTGTAGACACTCCCTGCCCTGACCCAAATACCTTGCACGAT	1200	
Qy	1201	GGCCGCGGCCACTGCCCCCTGCTGAGGCTCCCTCAACTGCTGCTCAATGGCTCCCTC	1260	
Db	1201	GGCCGCGGCCACTGCCCCCTGCTGAGGCTCCCTCAACTGCTGCTCAATGGCTCCCTC	1260	
Qy	1261	CACATCAGCAGCAGCCTGTGCCCCATGATCAGGGGTCCCTGGCCCGGGCCCTGTGGC	1320	
Db	1261	CACATCAGCAGCAGCCTGTGCCCCATGATCAGGGGTCCCTGGCCCGGGCCCTGTGGC	1320	
Qy	1321	TGCTGTCTCAGCTGCTGAACATTTGGAGCAAGGAAAGTCTCTCTACTGCTTGAGCCC	1380	
Db	1321	TGCTGTCTCAGCTGCTGAACATTTGGAGCAAGGAAAGTCTCTCTACTGCTTGAGCCC	1380	
Qy	1381	ACACTCTCAAAAAGGCCAACAAAAACAGACTTCAAGTCGGAGGTCATGTGTACAGTT	1440	
Db	1381	ACACTCTCAAAAAGGCCAACAAAAACAGACTTCAAGTCGGAGGTCATGTGTACAGTT	1440	
Qy	1441	AAAGCCAGTGGCAGCTCCACCTGCCCGGAGGCGACACTGGAGTCTGTCTCATTTATGTA	1500	
Db	1441	AAAGCCAGTGGCAGCTCCACCTGCCCGGAGGCGACACTGGAGTCTGTCTCATTTATGTA	1500	
Qy	1501	TTCTCTGGAGCTGTGCTATGTACCTGCAATTTTCAAGGAGACATATTCAACTCTCTGCT	1560	
Db	1501	TTCTCTGGAGCTGTGCTATGTACCTGCAATTTTCAAGGAGACATATTCAACTCTCTGCT	1560	
Qy	1561	CTTTCCAAAACCTGTGTCTATCTCCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGCTT	1620	
Db	1561	CTTTCCAAAACCTGTGTCTATCTCCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGCTT	1620	
Qy	1621	AGCAGAGCAGTTCCTGGCCAGAGAACACAGCTGCTGCCAGGGGCGAGCCCTCATAG	1680	
Db	1621	AGCAGAGCAGTTCCTGGCCAGAGAACACAGCTGCTGCCAGGGGCGAGCCCTCATAG	1680	
Qy	1681	GAGGCCACAGGAGGAGGCCCAAGCGCTAGAAGCTTTGTTGAAGCTGTGAGCAGGAGAAC	1740	
Db	1681	GAGGCCACAGGAGGAGGCCCAAGCGCTAGAAGCTTTGTTGAAGCTGTGAGCAGGAGAAC	1740	
Qy	1741	GGTGCCACACAGCTTCCAGGCTCTCCCTGACCTGCCCTGCTATGCCCCACACCCCTACG	1800	
Db	1741	GGTGCCACACAGCTTCCAGGCTCTCCCTGACCTGCCCTGCTATGCCCCACACCCCTACG	1800	

```
Qy 1801 CCGTGGCTCTGTCGACGTGTACGTAGATAGCTCTCGCCTGGGTCTGTGCTGTGTGCTGTA 1860
|||||
Db 1801 CCGTGGCTCTGTCGACGTGTACGTAGATAGCTCTCGCCTGGGTCTGTGCTGTGTGCTGTA 1860
|||||
Qy 1861 AAAGCTTAATGGGTGGCCAGGTGTGTACCTTCTCCCAAGCAAGCAATATGGAGCATC 1920
|||||
Db 1861 AAAGCTTAATGGGTGGCCAGGTGTGTACCTTCTCCCAAGCAAGCAATATGGAGCATC 1920
|||||
Qy 1921 TACCAGACGCCACTCTGCACACACTCACTCCACCTCTCAAGCCTCCAACTCTTGCC 1980
|||||
Db 1921 TACCAGACGCCACTCTGCACACACTCACTCCACCTCTCAAGCCTCCAACTCTTGCC 1980
|||||
Qy 1981 CAGATTGGGCTCATTAATGTCGTTGCTGCGCATCTGCATGATGACAGCAGCTCCCA 2040
|||||
Db 1981 CAGATTGGGCTCATTAATGTCGTTGCTGCGCATCTGCATGATGACAGCAGCTCCCA 2040
|||||
Qy 2041 TGGTGGTCTGCTGTGAGCTCTTCAAGTCTTAATCTTAACCTCAGGATTAGTCCCAAG 2100
|||||
Db 2041 TGGTGGTCTGCTGTGAGCTCTTCAAGTCTTAATCTTAACCTCAGGATTAGTCCCAAG 2100
|||||
Qy 2101 TGGCGTGAGACCCAGCAGCAGCTTCTGGCCCTCTCCCTGCTCAATCTAAAGCAGT 2160
|||||
Db 2101 TGGCGTGAGACCCAGCAGCAGCTTCTGGCCCTCTCCCTGCTCAATCTAAAGCAGT 2160
|||||
Qy 2161 GCCACCCCTCAAGTGAATAGAAAGAGTTCATGAGTGAAGGCTGCAAGCAATCTT 2220
|||||
Db 2161 GCCACCCCTCAAGTGAATAGAAAGAGTTCATGAGTGAAGGCTGCAAGCAATCTT 2220
|||||
Qy 2221 ATCTGGCCACATCTCTCGTGACACACACCCCAATGAGTTAACTTGAAGTTGACTAT 2280
|||||
Db 2221 ATCTGGCCACATCTCTCGTGACACACACCCCAATGAGTTAACTTGAAGTTGACTAT 2280
|||||
Qy 2281 TTTAATGTCGCCAGGAGTCTAATCTGCTCTGTCCCTCTGTTCCCTTTCTCTGTAAGTCC 2340
|||||
Db 2281 TTTAATGTCGCCAGGAGTCTAATCTGCTCTGTCCCTCTGTTCCCTTTCTCTGTAAGTCC 2340
|||||
Qy 2341 AGCACACATCTTGTCTTCCCTCCCGAGTTCTCTGCGCTCTCCACCCCTCCAGCTTCATGCTC 2400
|||||
Db 2341 AGCACACATCTTGTCTTCCCTCCCGAGTTCTCTGCGCTCTCCACCCCTCCAGCTTCATGCTC 2400
|||||
Qy 2401 AGTGTGTGCTTAATAAATGGACATATTTTCTCTTAAAAAAAAA 2447
|||||
Db 2401 AGTGTGTGCTTAATAAATGGACATATTTTCTCTTAAAAAAAAA 2447
|||||

RESULT 2
BC032787
LOCUS BC032787
DEFINITION Homo sapiens, calcium/calmodulin-dependent protein kinase IG, clone
MGC:44894 IMAGE:5179957, mRNA, complete cds.
ACCESSION BC032787
VERSION BC032787.1 GI:21619664
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2474)
Strausberg, R.
Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapb-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
```

Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masliello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 68 Row: K Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14196444.

FEATURES

source

1..2474
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:44894 IMAGE:5179957"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
70..1500
/codon_start=1
/product="calcium/calmodulin-dependent protein kinase IG"
/protein_id="AAH32787.1"
/db_xref="GI:21619664"
/translation="MGRKSDDCSWKQKTNIRKTFIPEVLGSCAFSEFLVKORL
TGKLFALCKIKSPAPRDSLENIAYLAKIKHENIVLTEDIYESTHYLVQVLVSG
GELDFRLERGVYETKDALVIOQVLSAVKYDHENGIVHRDLKPNLLYLTPTPENSKI
MITDFGLSRMEQINGTACGTGYVAPVLAQPKYSKAVDCMGVITVYILLCCYPP
FYDETSLKFEIKEGYEFESPFWDIISAKDFICHLEKLDPNRYTCERAKSHPW
IDGNTALHRDIYPSVSLQIKNFKAQFNAAVVHMKRLHMLHSPGVRPEVE
NRPEYQASSETSPSSPEITITAPVLDHVALPALTPCOHGRRTAPGGRSLNCL
VNSLHISLSLVPMHQGSANGCGCCSCLNIGSKGSKSYSEPTLLKKANKQNFK
SEWVPVKAGSSHCRAGTGVCLIM"

CDS

BASE COUNT 618 a 707 c 603 g 546 t
ORIGIN
Query Match 97.9%; Score 2396; DB 9; Length 2474;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2446; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 TGGAGTGGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 60
|||||
Db 1 TGGAGTGGGAGCTCAAGCAGGATTCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 60
|||||
Qy 61 CTGGAGGCAATGGTCCGAAGGAAGATGACTGCAGTTCCTCGGAAGAACAGACCACC 120
|||||
Db 61 CTGGAGGCAATGGTCCGAAGGAAGATGACTGCAGTTCCTCGGAAGAACAGACCACC 120
|||||
Qy 121 AACATCCGGAAGAACTTCATTTTATGAAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180
|||||
Db 121 AACATCCGGAAGAACTTCATTTTATGAAAGTCTGGGATCAGGAGCTTTCTCAGAAGTT 180
|||||
Qy 181 TTCCTGTTGAAGCAAGACTGACTGGGAAGCTCTTTCCTCTGAAGTGCATCAAGAAGTCA 240
|||||
Db 181 TTCCTGTTGAAGCAAGACTGACTGGGAAGCTCTTTCCTCTGAAGTGCATCAAGAAGTCA 240
|||||
Qy 241 CCTGCTTCCGGGACAGCAGCCTGGGAATGAGATTGCTGTGTGTAAGAAAGATCAAGCAT 300
|||||
Db 241 CCTGCTTCCGGGACAGCAGCCTGGGAATGAGATTGCTGTGTGTAAGAAAGATCAAGCAT 300
|||||
Qy 301 GAAACATTTGACCCCTGGAGGACATCTATGAGAGCACCACCCACTACTACTACCTGGTCATG 360
|||||
Db 301 GAAACATTTGACCCCTGGAGGACATCTATGAGAGCACCACCCACTACTACTACCTGGTCATG 360
|||||
Qy 361 CAGCTTGTCTGTTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAG 420
|||||

Db 361 CAGCTTGTCTGGTGGGAGCTTTTGACCGGATCCTGGAGGGGGTGTCTACACAGAG 420
Qy 421 AAGGATGCCAGTCTGGRGATCCAGAGGCTTGTTCGGCAGTGAATACCTACATGAGAT 480
Db 421 AAGGATGCCAGTCTGGRGATCCAGAGGCTTGTTCGGCAGTGAATACCTACATGAGAT 480
Qy 481 GGCATCTCCACAGAGACTTAAAGCCCGGAAACCTGCTTACCTTACCCTTGAAGAGAC 540
Db 481 GGCATCTCCACAGAGACTTAAAGCCCGGAAACCTGCTTACCTTACCCTTGAAGAGAC 540
Qy 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCC 600
Db 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCC 600
Qy 501 ACTGCCCTGGGAGCCCGAGCTACGTGGCTCCAGAGTGGTGGCCCGAGAAACCTTACAGC 560
Db 601 ACTGCCCTGGGAGCCCGAGCTACGTGGCTCCAGAGTGGTGGCCCGAGAAACCTTACAGC 560
Qy 661 AAGGCTGTGGATTGCTGGTCCATCGGCTCATCACCTACATATTGCTCTGTGGATACCC 720
Db 661 AAGGCTGTGGATTGCTGGTCCATCGGCTCATCACCTACATATTGCTCTGTGGATACCC 720
Qy 721 CCCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAG 780
Db 721 CCCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGAGGGCTACTATGAG 780
Qy 781 TTTGAGTCTCCATCTGGGATGACATTTCTGAGTTCAGCCAGGACTTTATTGGCCACTTG 840
Db 781 TTTGAGTCTCCATCTGGGATGACATTTCTGAGTTCAGCCAGGACTTTATTGGCCACTTG 840
Qy 841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCCCTTGAGTCAATCCCTGGATT 900
Db 841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCCCTTGAGTCAATCCCTGGATT 900
Qy 901 GACGGAACAGCGCCCTCCACGGGACATCTACCCATCATGTCAGCTCCAGATCCAGAG 960
Db 901 GACGGAACAGCGCCCTCCACGGGACATCTACCCATCATGTCAGCTCCAGATCCAGAG 960
Qy 961 AACTTTGCTTAAGACAAAGTGGAGCGCAAGCCTTCAACGCGAGCAGCTGTGGTGCACCATG 1020
Db 961 AACTTTGCTTAAGACAAAGTGGAGCAAGCCTTCAACGCGAGCAGCTGTGGTGCACCATG 1020
Qy 1021 AGGAAGTACACATGACATTCGACAGCCCGGGCGTCCGCCAGAGTGGAGAACAGCGCG 1080
Db 1021 AGGAAGTACACATGACATTCGACAGCCCGGGCGTCCGCCAGAGTGGAGAACAGCGCG 1080
Qy 1081 CCTGAAACTCAAGCCTCAGAAACCTCTAGACCAGCTCCCTGAGATCACCATCCCGAG 1140
Db 1081 CCTGAAACTCAAGCCTCAGAAACCTCTAGACCAGCTCCCTGAGATCACCATCCCGAG 1140
Qy 1141 GCACCTGTCTGGACACAGTGTAGCACTCCCTGGCCCTGACCCAAATTACCCCTGCCAGAT 1200
Db 1141 GCACCTGTCTGGACACAGTGTAGCACTCCCTGGCCCTGACCCAAATTACCCCTGCCAGAT 1200
Qy 1201 GGCCGCGGCCCACTGCCCTGGTGGCAGTCCCTCAACTGCTGGTCAATGGCTCCCTC 1260
Db 1201 GGCCGCGGCCCACTGCCCTGGTGGCAGTCCCTCAACTGCTGGTCAATGGCTCCCTC 1260
Qy 1261 CACATCAGCAGCAGCCCTGGTGGCCATGCATCAGGGGTCCCTGGCCGCGGCCCTGTGGC 1320
Db 1261 CACATCAGCAGCAGCCCTGGTGGCCATGCATCAGGGGTCCCTGGCCGCGGCCCTGTGGC 1320
Qy 1321 TGCCTGCTCCAGCTGCCCTGAACATTTGGAGCAAAAGAAAGTCCCTACTGCTTGAGCC 1380
Db 1321 TGCCTGCTCCAGCTGCCCTGAACATTTGGAGCAAAAGAAAGTCCCTACTGCTTGAGCC 1380
Qy 1381 ACACCTCTCAAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCAATGGTACCAGTT 1440
Db 1381 ACACCTCTCAAAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTCAATGGTACCAGTT 1440
Qy 1441 AAAGCCAGTGGCAGCTCCCACTGCCGGGAGGCGAGACTGGAGTCTGTCTCATTTATGTGA 1500
Db 1441 AAAGCCAGTGGCAGCTCCCACTGCCGGGAGGCGAGACTGGAGTCTGTCTCATTTATGTGA 1500

Qy 1501 TTCTGGAGCCCTGGCTATGTCACTGCAATTTTCAGAGACATATTTCAACTCCTCTGCT 1560
Db 1501 TTCTGGAGCCCTGGCTATGTCACTGCAATTTTCAGAGACATATTTCAACTCCTCTGCT 1560
Qy 1561 CTTTCAAAACCTTGGTGTCTATCCGGCAGAGGGAGGAGCAGCAAGTGGAGCAGGCGTT 1620
Db 1561 CTTTCAAAACCTTGGTGTCTATCCGGCAGAGGGAGGAGCAGCAAGTGGAGCAGGCGTT 1620
Qy 1621 AGCAGGACAGTTTCTTGCCAGAGACCCAGCCTGCTGCCAGCGGGGAGCCCTCATAG 1680
Db 1621 AGCAGGACAGTTTCTTGCCAGAGACCCAGCCTGCTGCCAGCGGGGAGCCCTCATAG 1680
Qy 1681 GAGGCCAGGAGGGAGCCCGGAGGCTAGAGCCCTTGTGAAGCTGTGAGCAGGAGAGC 1740
Db 1681 GAGGCCAGGAGGGAGCCCGGAGGCTAGAGCCCTTGTGAAGCTGTGAGCAGGAGAGC 1740
Qy 1741 GGTGCCACACAGCTTCCAGGCTCTCCCTGACCTGCTCTATGCCCCACACCCCTACGTG 1800
Db 1741 GGTGCCACACAGCTTCCAGGCTCTCCCTGACCTGCTCTATGCCCCACACCCCTACGTG 1800
Qy 1801 CCGTGGCTCTGTGAGTGTACGTAGATAGCTCTGCCTGGGTCTGTGCTTGTCTGTGA 1860
Db 1801 CCGTGGCTCTGTGAGTGTACGTAGATAGCTCTGCCTGGGTCTGTGCTTGTCTGTGA 1860
Qy 1861 AAAGCTTAATGGCTGGCAGGCTGTGTACCTTCTCCAAAGCAAGCATATGGAGCATC 1920
Db 1861 AAAGCTTAATGGCTGGCAGGCTGTGTACCTTCTCCAAAGCAAGCATATGGAGCATC 1920
Qy 1921 TACCCAGACTCCCACTGTGCACACACTCCTCCACCTCTCAAGCCTCCCAACCTTGGC 1980
Db 1921 TACCCAGACTCCCACTGTGCACACACTCCTCCACCTCTCAAGCCTCCCAACCTTGGC 1980
Qy 1981 CAGATTGGGCTCATTAATGTGCTTGGCTGCCCATCTGCATGAATGACAGGAGCTCCCA 2040
Db 1981 CAGATTGGGCTCATTAATGTGCTTGGCTGCCCATCTGCATGAATGACAGGAGCTCCCA 2040
Qy 2041 TGGTGGTCTGCTGTGAGCTCTTCAAGTTCTTAATCCTTAATCCTCAGGATTAAGCTCCCAAG 2100
Db 2041 TGGTGGTCTGCTGTGAGCTCTTCAAGTTCTTAATCCTTAATCCTCAGGATTAAGCTCCCAAG 2100
Qy 2101 TGCCTGAGAGCCAGCCAGCAGCAGCTTCTGCGCTTCTCCCTGCCTCAATCTAAGAGCAGT 2160
Db 2101 TGCCTGAGAGCCAGCCAGCAGCAGCTTCTGCGCTTCTCCCTGCCTCAATCTAAGAGCAGT 2160
Qy 2161 GCCACACCTCCAAAGTGGAAATAGAAAGATTCATGATGAAGGGCTGCAAGGAATCTT 2220
Db 2161 GCCACACCTCCAAAGTGGAAATAGAAAGATTCATGATGAAGGGCTGCAAGGAATCTT 2220
Qy 2221 ATCTTGGCCACATGTCTCCGTGCACACACCCCAATGGAGTTAACTTGGAGTTGACTAT 2280
Db 2221 ATCTTGGCCACATGTCTCCGTGCACACACCCCAATGGAGTTAACTTGGAGTTGACTAT 2280
Qy 2281 TTTAATGTCTGCCAGGAGTTCTAATCCTGCTCTGCTTCCCTTTCTCTCTCTTGAAGTCC 2340
Db 2281 TTTAATGTCTGCCAGGAGTTCTAATCCTGCTCTGCTTCCCTTTCTCTCTCTTGAAGTCC 2340
Qy 2341 AGCAGACCATCTTGTGCTTCCCAAGTTTCTGCGCTTCCACCCCTCCAGCTTCATGCTC 2400
Db 2341 AGCAGACCATCTTGTGCTTCCCAAGTTTCTGCGCTTCCACCCCTCCAGCTTCATGCTC 2400
Qy 2401 AGTGTGTGCTTAATAAATGAGCATATTTTCTCTAAAAA 2447
Db 2401 AGTGTGTGCTTAATAAATGAGCATATTTTCTCTAAAAA 2447

RESULT 3

AK095713

LOCUS

DEFINITION

ACCESSION

VERSION

AK095713 Homo sapiens cDNA FLJ38394 fis, clone FEBRA2007534, moderately similar to CamK1-like protein kinase.
AK095713
AK095713.1 GI:21755035

2612 bp mRNA linear PRI 15-JUL-2002

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
Source
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches 2393; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 39 GGCATCCTCAGAGCTTCAACTCTGGAGGCAATGGTGCAGAAAGGAAGATGACTGCAG 98
Db 215 GGCATCCTCAGAGCTTCAACTCTGGAGGCAATGGTGCAGAAAGGAAGATGACTGCAG 274
QY 99 TTCCTGGAAGAACACACACCAACATCGGAAACCTTCATTTTATGGAAGTGCCTGGG 158
Db 275 TTCCTGGAAGAACACACCAACATCGGAAACCTTCATTTTATGGAAGTGCCTGGG 334
QY 159 ATCAGGAGCTTCTCAGAAAGTTTCTGCTGGAAGCAAAAGCTGACTGGGAAGCTCTTTGC 218
Db 335 ATCAGGAGCTTCTCAGAAAGTTTCTGCTGGAAGCAAAAGCTGACTGGGAAGCTCTTTGC 394
QY 219 TCTGGAAGTGCATCAAGAGTCACTGCTTCCGGACACAGCCCTGGAGAATGAGATTGC 278
Db 395 TCTGGAAGTGCATCAAGAGTCACTGCTTCCGGACACAGCCCTGGAGAATGAGATTGC 454
QY 279 TGTGTTGAAAAGATCAAGCATGAAACATTTGTGACCTTGGAGGACATCTATGAGAGCAC 338
Db 455 TGTGTTGAAAAGATCAAGCATGAAACATTTGTGACCTTGGAGGACATCTATGAGAGCAC 514
QY 339 CACCCACTACTACTGGTGCATGAGCTTCTTCTGGTGGGAGCTCTTTGACCGGATCCT 398
Db 515 CACCCACTACTACTGGTGCATGAGCTTCTTCTGGTGGGAGCTCTTTGACCGGATCCT 574
QY 399 GGAGCGGGTGCTACACAGAGAGGATCCAGTCTGGTGATCCAGAGGCTTGTCTCGGC 458
|||||

oligo capping; fls (full insert sequence).
Homo sapiens fetal brain cDNA to mRNA, clone_lib:FEBRA2
clone:FEBRA2007534.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M.,
Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yanashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2612)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1. .2612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FEBRA2007534"
/tissue_type="brain"
/clone_lib="FEBRA2"
/dev_stage="fetal"
/note="cloning vector: pME18SFL3"
BASE COUNT 619 a 750 c 645 g 598 t

Db 575 GGAGCGGGTGCTACACAGAGAAGGATGCCAGTCTGTGTATCCAGAGGCTTTGTCCGC 634
QY 459 AGTGAATACCTACATGAGAATGGCATGCTCCACAGAGACTTTAAAGCCCGGAAACCTGCT 518
Db 635 AGTGAATACCTACATGAGAATGGCATGCTCCACAGAGACTTTAAAGCCCGGAAACCTGCT 694
QY 519 TTACCTTACCCCTGAAGAGAAGCTCTAAGATCATGATCACTGACTTTGGTCTGTCTCAAGAT 578
Db 695 TTACCTTACCCCTGAAGAGAAGCTCTAAGATCATGATCACTGACTTTGGTCTGTCTCAAGAT 754
QY 579 GGAACAGAATGSCATCATGCTCCACTGCTGGACCCAGGCTACGTTGGCTCCAGAGT 638
Db 755 GGAACAGAATGSCATCATGCTCCACTGCTGGACCCAGGCTACGTTGGCTCCAGAGT 814
QY 639 GCTGGCCCGAGAAACCTTACAGCAAGGCTGTGATTTGCTGCTCCATCGGCGTCATCACCTA 698
Db 815 GCTGGCCCGAGAAACCTTACAGCAAGGCTGTGATTTGCTGCTCCATCGGCGTCATCACCTA 874
QY 699 CATATTGCTCTGTGGATACCCCGCTTCTATGAAGAAAGGAGTCTAAGCTTTTCGAGAA 758
Db 875 CATATTGCTCTGTGGATACCCCGCTTCTATGAAGAAAGGAGTCTAAGCTTTTCGAGAA 934
QY 759 GATCAAGGAGGCTACTATGATTTGAGTCTCCATTCTGGGATGACATTTCTGAGTCAAG 818
Db 935 GATCAAGGAGGCTACTATGATTTGAGTCTCCATTCTGGGATGACATTTCTGAGTCAAG 994
QY 819 CAAGGACTTTATTGCGCACTTCTGTTGAGAAGGATCCGAACGAGCGGTACACTGTGAGAA 878
Db 995 CAAGGACTTTATTGCGCACTTCTGTTGAGAAGGATCCGAACGAGCGGTACACTGTGAGAA 1054
QY 879 GGCTTTGAGTCACTCCCTGGATTGACGGAACACGCGCTCCACCGGACATCTACCCATC 938
Db 1055 GGCTTTGAGTCACTCCCTGGATTGACGGAACACGCGCTCCACCGGACATCTACCCATC 1114
QY 939 AGTCAGCTCCAGATCCAGAAAGCTTTGCTAAGAGCAAGTGGAGGCAAGCTTCAACGC 998
Db 1115 AGTCAGCTCCAGATCCAGAAAGCTTTGCTAAGAGCAAGTGGAGGCAAGCTTCAACGC 1174
QY 999 AGCAGCTGTGGTGCACCATGAGGAAGCTACACATGACCTGACAGCGCGGGGCTCGG 1058
Db 1175 AGCAGCTGTGGTGCACCATGAGGAAGCTACACATGAACTGACAGCGCGGGGCTCGG 1234
QY 1059 CCAGAGGTGGAGAACAGCGCGCTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTC 1118
Db 1235 CCAGAGGTGGAGAACAGCGCGCTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTC 1294
QY 1119 CCCTGAGATCAACATCAGGAGCACTTCTCTGGACACAGTGTAGCACTCCCTGCCCT 1178
Db 1295 CCCTGAGATCAACATCAGGAGCACTTCTCTGGACACAGTGTAGCACTCCCTGCCCT 1354
QY 1179 GACCCAAATTACCTGCCAGATGGCGCGCGCCCTGCTGGCAGTCCCTCAA 1238
Db 1355 GACCCAAATTACCTGCCAGATGGCGCGCGCCCTGCTGGCAGTCCCTCAA 1414
QY 1239 CTGCTGTGTAATPGGTCCCTCCATCAGCAGCAGCCTGTGGTCCCATGATCAGGGGTC 1298
Db 1415 CTGCTGTGTAATPGGTCCCTCCATCAGCAGCAGCCTGTGGTCCCATGATCAGGGGTC 1474
QY 1299 CTTGGCGCGCGGGCCCTGTGGTCTGCTGCTCCAGTTCCTTGGAGCAAAAGGAAA 1358
Db 1475 CTTGGCGCGCGGGCCCTGTGGTCTGCTGCTCCAGTTCCTTGGAGCAAAAGGAAA 1534
QY 1359 GTCTCTTACTGCTCTGAGCCCACTCTCTCAAAAAGGCCAACAAAAACAGAACTCAA 1418
Db 1535 GTCTCTTACTGCTCTGAGCCCACTCTCTCAAAAAGGCCAACAAAAACAGAACTCAA 1594
QY 1419 GTGGAGGTCATGGTACCAGTTAAAGCCAGTGGCAGCTCCACTGCCGGGAGGCGAGAC 1478
Db 1595 GTGGAGGTCATGGTACCAGTTAAAGCCAGTGGCAGCTCCACTGCCGGGAGGCGAGAC 1654
QY 1479 TGGAGTCTGTCTCATTTATGATTCTCTGGAGCCTGTGGTCTATGCTCAATTTTCAGG 1538
Db 1655 TGGAGTCTGTCTCATTTATGATTCTCTGGAGCCTGTGGTCTATGCTCAATTTTCAGG 1714
|||||

QY	1539	AGACATATTCAACTCTCTGCTCTTCCAAACCTGGTGTATATCCGGCAGAGGGAAGG	1598
Db	1715	AGACATATTCAACTCTCTGCTCTTCCAAACCTGGTGTATATCCGGCAGAGGGAAGG	1774
QY	1599	CAGAGCAAGTGGAGCAGGCTTAGCAGGAGCAGTTTCTGGCCAGAGCACCAGCCTGCTG	1658
Db	1775	CAGAGCAAGTGGAGCAGGCTTAGCAGGAGCAGTTTCTGGCCAGAGCACCAGCCTGCTG	1834
QY	1659	CCAGCGGGGAGCCCTCATAGAGGCCCAGAGGAGGAGCCCAAGCGGTAGAACCTTGTGT	1718
Db	1835	CCAGCGGGGAGCCCTCATAGAGGCCCAGAGGAGGAGCCCAAGCGGTAGAACCTTGTGT	1894
QY	1719	TGAAGCTGTGAGCAGAGAGGCGTCCACCAGCTTCCAGCTTCCCTGAGCTCCCTGC	1778
Db	1895	TGAAGCTGTGAGCAGAGAGGCGTCCACCAGCTTCCAGCTTCCCTGAGCTCCCTGC	1954
QY	1779	TCTATGCCCCACACCTACGTGCGGTCTGTGTCAGTGTACGTAGATAGCTCTCGCT	1838
Db	1955	TCTATGCCCCACACCTACGTGCGGTCTGTGTCAGTGTACGTAGATAGCTCTCGCT	2014
QY	1839	GGTCTGTGCTGTTTCTGTGTAAGCTTAATGGGCTGGCCAGGCTGTGTCACTTCTCC	1898
Db	2015	GGTCTGTGCTGTTTCTGTGTAAGCTTAATGGGCTGGCCAGGCTGTGTCACTTCTCC	2074
QY	1899	AAGCAAGGCCATATGAGCATCTACCCAGCTCCCACTCTGCACACACTCACTCCCACT	1958
Db	2075	AAGCAAGGCCATATGAGCATCTACCCAGCTCCCACTCTGCACACACTCACTCCCACT	2134
QY	1959	CTCAAGCCTCCAACTCTTGGCCAGATTGGGCTCATTAATGTGTTGGCTGCCATCTGC	2018
Db	2135	CTCAAGCCTCCAACTCTTGGCCAGATTGGGCTCATTAATGTGTTGGCTGCCATCTGC	2194
QY	2019	ATGAATGACAGGAGCTCCCATGTTGGTCTGCCTGTGAGCTTCTCAAGTTCTAATCCTT	2078
Db	2195	ATGAATGACAGGAGCTCCCATGTTGGTCTGCCTGTGAGCTTCTCAAGTTCTAATCCTT	2254
QY	2079	AACTCCAGGATTAGCTCCCAAGTGCCTGAGACCCAGCCAGCACACTTCTGGCCCTTCTC	2138
Db	2255	AACTCCAGGATTAGCTCCCAAGTGCCTGAGACCCAGCCAGCACACTTCTGGCCCTTCTC	2314
QY	2139	CTGCGCTCAATCTAAAGCAGTGCCACACCCCTCCAAAGTGAATAGAAAGAGTTCATGA	2198
Db	2315	CTGCGCTCAATCTAAAGCAGTGCCACACCCCTCCAAAGTGAATAGAAAGAGTTCATGA	2374
QY	2199	GTAAAGGCTGCAAGGAATCTTATCTGCGCACATCTCTCGTGCACACACCCCAATGA	2258
Db	2375	GTAAAGGCTGCAAGGAATCTTATCTGCGCACATCTCTCGTGCACACACCCCAATGA	2434
QY	2259	GTTAACTTGGAAAGTGAATTTTAAATGCTGCGCAGGAGTTCTAATCCTGCTCTGTTCT	2318
Db	2435	GTTAACTTGGAAAGTGAATTTTAAATGCTGCGCAGGAGTTCTAATCCTGCTCTGTTCT	2494
QY	2319	CTTTTCTCTCCTTGAAGTCCAGCACCATCTTGTCTTCCCTCCAGTTTCTCGCCCT	2378
Db	2495	CTTTTCTCTCCTTGAAGTCCAGCACCATCTTGTCTTCCCTCCAGTTTCTCGCCCT	2554
QY	2379	CCACCCCTCAGCTTCTGCTAGTGTGTGCTTAATAAATGGACATATTTTCTCT	2436
Db	2555	CCACCCCTCAGCTTCTGCTAGTGTGTGCTTAATAAATGGACATATTTTCTCT	2612
RESULT 4			
AF428261			
LOCUS			
DEFINITION			
Homo sapiens calcium/calmodulin-dependent protein kinase I gamma			
(CAMK1G) mRNA, complete cds.			
ACCESSION			
AF428261			
VERSION			
AF428261.1 GI:16755791			
KEYWORDS			
Homo sapiens.			
SOURCE			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2464)		
TITLE	Schutte,B.C., Bjork,B.C., Coppage,K.B., Malik,M.I., Gregory,S.G., Scott,D.J., Brentzelli,L.M., Watanabe,Y., Dixon,M.J. and Murray,J.C. A preliminary gene map for the Van der Woude syndrome critical region derived from 900 kb of genomic sequence at 1q32-q41		
JOURNAL	Genome Res. 10 (1), 81-94 (2000)		
MEDLINE	20113118		
PUBMED	10645953		
REFERENCE	2 (bases 1 to 2464)		
AUTHORS	Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.		
TITLE	Characterization of the human ortholog of rat Cam Kinase I gamma (CamK1g) at 1q32-q41		
JOURNAL	Unpublished		
AUTHORS	3 (bases 1 to 2464)		
TITLE	Bjork,B.C., Watanabe,Y., Murray,J.C. and Schutte,B.C.		
JOURNAL	Direct Submission		
FEATURES	Submitted (05-OCT-2001) Genetics, University of Iowa, 140 EMRB, Iowa City, IA 52242, USA		
Source	Location/Qualifiers		
gene	1..2464		
CDS	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="1q32-q41"		
	1..2464		
	/gene="CAMK1G"		
	73..1503		
	/gene="CAMK1G"		
	/codon_start=1		
	/product="calcium/calmodulin-dependent protein kinase I gamma"		
	/protein_id="AAL28100.1"		
	/db_xref="GI:16755792"		
	/translation="MGRBEDDCSSWKKTNRKTFIFMEVLGSGAFSEFLVKQRL TGKLFALKIKKSPAFRDSLENEIAVLKKIKHENIVLEIDYESTHYLLVMQLVSG GELFDRLIERGVYTERKDSLVIQQVLSAVKYLHENGIVHRDLKPNLLYLTPEENSKI MYDFGLSKMEQNGINSTAGTGPYVAPELAQKPYSKADWCSTGIVTYILLGYPF FYBETESKLFEIKEGYEFESPWDDIISAKDFICHLEKDPNERVTCERKALSHPW IDGNTALHRDIYPSVSLOIQKNFAKSKRWQAFNAAVVHMRKLMHNSPGIRPEVE NRPPEQASSETSRPSPETITIEAPVLDHVALPALTOIPCOHGRPTAPGGRSLNCL VNGSLHSSLSVPMHOGSLAAGPCGCCSCLNIGNKGSYSYSEPTLLKANKKONFK SEYVVPVKASGSGSHCRAGDTGVCLIM"		
polyA_signal	2415..2420		
BASE COUNT	609 a	708 c	602 g 545 t
ORIGIN	/gene="CAMK1G"		
Query Match	83.7%	Score 2048;	DB 9; Length 2464;
Best Local Similarity	99.8%	Pred. No. 0;	
Matches 2418;	Conservative	0; Mismatches	5; Indels 1; Gaps 1;
QY	24	TCTTCCCGAGTCCCTGGCATCCTCAGAACTTCAACTCTGAGGCAATGGTGCAGAGA	83
Db	27	TCTTCCCGAGTCCCTGGCATCCTCAGAACTTCAACTCTGAGGCAATGGTGCAGAGA	86
QY	84	AGAAGATGACTGCTAGTTCCTGGAGAGAACAGACACCAACATCCGGAACCTTCATTT	143
Db	87	AGAAGATGACTGCTAGTTCCTGGAGAGAACAGACACCAACATCCGGAACCTTCATTT	146
QY	144	TATGGAAGTCTGGGATCAGGAGCTTCTCAGAACTTTCCTGTTGAAGCAAGACTGAC	203
Db	147	TATGGAAGTCTGGGATCAGGAGCTTCTCAGAACTTTCCTGTTGAAGCAAGACTGAC	206
QY	204	TGGGAAGCTCTTCTCTGAAGTGCATCAAGAAGTCACTACCTTCCTCGGACAGCAGCT	263
Db	207	TGGGAAGCTCTTCTCTGAAGTGCATCAAGAAGTCACTACCTTCCTCGGACAGCAGCT	266
QY	264	GGAAATGAGATTCCTGTGTGAAAAAGATCAACAGCATGAAACATTTGACCCCTGGAGA	323
Db	267	GGAAATGAGATTCCTGTGTGAAAAAGATCAACAGCATGAAACATTTGACCCCTGGAGA	326
QY	324	CATCTATGAGAGCACCACCCACTACTACCTGGTCTATGAGCTTTTCTCTGTTGGGAGCT	383

|||||
Db 327 CATCTATGAGACACACCCACTACTACCTGGTCAATGACAGCTTGTCTGGTGGGAGCT 386
QY 384 CTTTGGACGGATCCTGGAGCGGGTGTCTACACAGAGAGGATGCCAGTCTGGTGTATCCA 443
Db 387 CTTTGGACGGATCCTGGAGCGGGTGTCTACACAGAGAGGATGCCAGTCTGGTGTATCCA 446
QY 444 GCAGGTCTTGTGGCAGTGAATACCTACATGAGAATGGCATGCTCCACAGAGACTTAAA 503
Db 447 GCAGGTCTTGTGGCAGTGAATACCTACATGAGAATGGCATGCTCCACAGAGACTTAAA 506
QY 504 GCCCGAAAACCTGTCTTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTT 563
Db 507 GCCCGAAAACCTGTCTTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTT 566
QY 564 TGGTCTGTCCAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 623
Db 567 TGGTCTGTCCAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 626
QY 624 CGTGGCTCCAGAAGTGTCTGGCCAGAACCCCTACAGCAAGGCTGTGGATGGTGGTCCAT 683
Db 627 CGTGGCTCCAGAAGTGTCTGGCCAGAACCCCTACAGCAAGGCTGTGGATGGTGGTCCAT 686
QY 684 CGCGGTCACTACATATCTCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTC 743
Db 687 CGCGGTCACTACATATCTCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTC 746
QY 744 TAAGCTTTTCGAGAGATCAAGAGGGCTACTATGAGTTTGAGTCTCCATCTCGGATGA 803
Db 747 TAAGCTTTTCGAGAGATCAAGAGGGCTACTATGAGTTTGAGTCTCCATCTCGGATGA 806
QY 804 CATTTCTGATCAGCACAGGACTTTATTTGGCATACCCCGCTTCTATGAAGAAACGGAGTC 863
Db 807 CATTTCTGATCAGCACAGGACTTTATTTGGCATACCCCGCTTCTATGAAGAAACGGAGTC 866
QY 864 GTACACCTGTGAGAAAGGCTTGTGATCATCCCTGGATTGACGAAACACAGCCCTCCACCG 923
Db 867 GTACACCTGTGAGAAAGGCTTGTGATCATCCCTGGATTGACGAAACACAGCCCTCCACCG 926
QY 924 GGNATCTACCATCAGTCAGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 983
Db 927 GGNATCTACCATCAGTCAGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 986
QY 984 GCAAGCTTTCAACGACGAGCTGTGGTGCACACATGAGGAGCTTACACATGAACCTGCA 1043
Db 987 GCAAGCTTTCAACGACGAGCTGTGGTGCACACATGAGGAGCTTACACATGAACCTGCA 1046
QY 1044 CAGCCGCGGCTCCGCCAGAGGTGGAGACAGCGCCGCTGAAACTCAAGCTCAGAAAC 1103
Db 1047 CAGCCGCGGCTCCGCCAGAGGTGGAGAACAGCGCCGCTGAAACTCAAGCTCAGAAAC 1106
QY 1104 CTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACCTGTCTTGGACACAGTGT 1163
Db 1107 CTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACCTGTCTTGGACACAGTGT 1166
QY 1164 AGCAGTCCCTGACCCCAATATCCCTGACAGATGGCGCGGCCCACTGCCCCCTGG 1223
Db 1167 AGCAGTCCCTGACCCCAATATCCCTGACAGATGGCGCGGCCCACTGCCCCCTGG 1226
QY 1224 TGGCAGGTCCCTCAACTGCTGTCATAGGCTCCCTCCACATCAGCAGCAGCTGGTGCC 1283
Db 1227 TGGCAGGTCCCTCAACTGCTGTCATAGGCTCCCTCCACATCAGCAGCAGCTGGTGCC 1286
QY 1284 CATGCAATCAGGGTCCCTGGCGCGGCCCTGTGGCTGCTGCTCCAGCTGCCTGAACAT 1343
Db 1287 CATGCAATCAGGGTCCCTGGCGCGGCCCTGTGGCTGCTGCTCCAGCTGCCTGAACAT 1346
QY 1344 TGGGAGCAAGGAAATCCTCTTACTGCTCTGAGGCCACACTCTCAAAAAGGCCAACAA 1403
Db 1347 TGGGAGCAAGGAAATCCTCTTACTGCTCTGAGGCCACACTCTCAAAAAGGCCAACAA 1406
QY 1404 AAAACAGAACTTCAAGTCGGAGTCAATGTACAGTTAAAGCCAGTGGCAGCTCCCACTG 1463
|||||

RESULT 5
HS272L161

Db 1407 AAAACAGAACTTCAAGTCGGAGGTCAATGATACCAAGTTAAAGCCAGTGGCAGCTCCCACTG 1466
QY 1464 CCGGGCAGGGCAGACTGGAGTCTGTCTCATATGATGATTTCTGGAGCCTGTGCCTATGTC 1523
Db 1467 CCGGGCAGGGCAGACTGGAGTCTGTCTCATATGATGATTTCTGGAGCCTGTGCCTATGTC 1526
QY 1524 ACTGCAATTTTCAGGAGACATATCAACTCTCTCTCTTCCAAACCTGTGTCTATCCG 1583
Db 1527 ACTGCAATTTTCAGGAGACATATCAACTCTCTCTCTTCCAAACCTGTGTCTATCCG 1586
QY 1584 GCAGAGGAGAGAGCAGAGCAAGTGGAGCGGCTTAGCAGGACAGTCTTGGCCAGA 1643
Db 1587 GCAGAGGAGAGAGCAGAGCAAGTGGAGCGGCTTAGCAGGACAGTCTTGGCCAGA 1646
QY 1644 AGCACCAGCTCTGCCAGCGGGCAGCCCTCATATAGAGGCCAGGAGGAGGCCCAAG 1703
Db 1647 AGCACCAGCTCTGCCAGCGGGCAGCCCTCATATAGAGGCCAGGAGGAGGCCCAAG 1706
QY 1704 GGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCT 1763
Db 1707 GGTAGAAGCCTTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCT 1766
QY 1764 CCCTGACCTGCTCTATGCCCCACACCCCTTACCTGCGCTGGCTCTGTGCAGTGTAGCT 1823
Db 1767 CCCTGACCTGCTCTATGCCCCACA - CCTACGTGCGCTGGCTCTGTGCAGTGTAGCT 1825
QY 1824 AGATAGCTCTGCGCTGGGTCTGTGCTGTTGCTGTGCTGAAAAGCTTAATGGCTGGCCAGGC 1883
Db 1826 AGATAGCTCTGCGCTGGGTCTGTGCTGTTGCTGTGAAAAGCTTAATGGCTGGCCAGGC 1885
QY 1884 TGTGTCACTTCTCCAAGCAAAAGCCATATGGAGCATCTACCCAGACTCCCACTCTGCACA 1943
Db 1886 TGTGTCACTTCTCCAAGCAAAAGCCATATGGAGCATCTACCCAGACTCCCACTCTGCACA 1945
QY 1944 CACTCACTCCCACTCTCAAGCCTCAAGCCTTGGCCAGATTTGGGCTCAATTAATGTGCT 2003
Db 1946 CACTCACTCCCACTCTCAAGCCTCAAGCCTTGGCCAGATTTGGGCTCAATTAATGTGCT 2005
QY 2004 TGCTCCCACTCTGCATGAATGACAGGAGCTCCCACTGCTGCTGTGAGCTCTT 2063
Db 2006 TGCTCCCACTCTGCATGAATGACAGGAGCTCCCACTGCTGCTGTGAGCTCTT 2065
QY 2064 CAAAGTTCTAATCCTTAACTCCAGGATTAAGTCCCAAGTGGCTGAGACCCAGCCAGACA 2123
Db 2066 CAAAGTTCTAATCCTTAACTCCAGGATTAAGTCCCAAGTGGCTGAGACCCAGCCAGACA 2125
QY 2124 CTTCTGGCCCTTCTCCCTGCTCAATCTAAAAGCAGTGCACACCTTCCAAAGTGAATA 2183
Db 2126 CTTCTGGCCCTTCTCCCTGCTCAATCTAAAAGCAGTGCACACCTTCCAAAGTGAATA 2185
QY 2184 GAAAGAGTTCTAGTAAGGCTGCAAGGAAATTTCTTATCTTGGCCACATGTCTCCGCTG 2243
Db 2186 GAAAGAGTTCTAGTAAGGCTGCAAGGAAATTTCTTATCTTGGCCACATGTCTCCGCTG 2245
QY 2244 CACACACCAATGAGGATTAACCTTGAAGTTGACTATTTTAAATGCTTGCAGGAGTTCTA 2303
Db 2246 CACACACCAATGAGGATTAACCTTGAAGTTGACTATTTTAAATGCTTGCAGGAGTTCTA 2305
QY 2304 ATCTGCTCTGTCTCCCTTCTCTCTTCTTGAAGTCCAGCACACCATTTCTTCTCTTCCC 2363
Db 2306 ATCTGCTCTGTCTCCCTTCTCTCTTCTTGAAGTCCAGCACACCATTTCTTCTCTCCC 2365
QY 2364 CAGTTTCTGCGCCTCCACCCCTCCAGCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2423
Db 2366 CAGTTTCTGCGCCTCCACCCCTCCAGCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2425
QY 2424 CATATTTTCTCTAAAAA 2447
Db 2426 CATATTTTCTCTAAAAA 2449
|||||

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence is the entire insert of clone 272L16. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

272L16 is from the library RPCI1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:pcrPAC2> This sequence was generated from a human chromosome 1 bacterial clone contig constructed in collaboration by the Sanger Centre chromosome 1 mapping group and Brian Schutte, Bryan Bjork, Kevin Coppage and Jeffrey Murray, Department of Pediatrics, University of Iowa, USA. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl>.

Location/Qualifiers
1. .157875
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="q32.1-32.3"
/clone="RPI-272L16"
/clone_lib="RPCI-1"
966. .1078
/note="MLT11 repeat: matches 281. .410 of consensus"
repeat_region
/note="MLT11 repeat: matches 281. .410 of consensus"
repeat_region
/note="MSRD repeat: matches 1. .426 of consensus"
repeat_region
/note="MLT1A1 repeat: matches 1. .319 of consensus"
repeat_region
/note="MLT11 repeat: matches 58. .203 of consensus"
repeat_region
/note="MER34 repeat: matches 6. .543 of consensus"
repeat_region
/note="MLT1A2 repeat: matches 23. .374 of consensus"
repeat_region
/note="MER54B repeat: matches 486. .902 of consensus"
repeat_region
/note="MLT1A2-internal repeat: matches 358. .1643 of consensus"
repeat_region
/note="AluJo repeat: matches 101. .302 of consensus"
repeat_region
/note="MLT1A1-internal repeat: matches 5. .450 of consensus"
repeat_region
/note="MLT1A1 repeat: matches 318. .365 of consensus"
repeat_region
/note="LNR7 repeat: matches 1. .450 of consensus"
repeat_region
/note="MLT1A1 repeat: matches 1. .318 of consensus"
repeat_region
/note="MIR repeat: matches 56. .142 of consensus"
repeat_region
/note="MIR repeat: matches 95. .146 of consensus"
repeat_region
/note="MIR repeat: matches 68. .212 of consensus"
repeat_region
/note="MIR repeat: matches 1. .711 of consensus"
repeat_region
/note="MER50 repeat: matches 13. .262 of consensus"
repeat_region
/note="MIR repeat: matches 13. .262 of consensus"
repeat_region
/note="Alusx repeat: matches 1. .304 of consensus"
repeat_region
/note="L2 repeat: matches 2506. .2695 of consensus"
repeat_region
10929. .11005

FEATURES
source
1. .157875
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="q32.1-32.3"
/clone="RPI-272L16"
/clone_lib="RPCI-1"
966. .1078
/note="MLT11 repeat: matches 281. .410 of consensus"
repeat_region
/note="MLT11 repeat: matches 281. .410 of consensus"
repeat_region
/note="MSRD repeat: matches 1. .426 of consensus"
repeat_region
/note="MLT1A1 repeat: matches 1. .319 of consensus"
repeat_region
/note="MLT11 repeat: matches 58. .203 of consensus"
repeat_region
/note="MER34 repeat: matches 6. .543 of consensus"
repeat_region
/note="MLT1A2 repeat: matches 23. .374 of consensus"
repeat_region
/note="MER54B repeat: matches 486. .902 of consensus"
repeat_region
/note="MLT1A2-internal repeat: matches 358. .1643 of consensus"
repeat_region
/note="AluJo repeat: matches 101. .302 of consensus"
repeat_region
/note="MLT1A1-internal repeat: matches 5. .450 of consensus"
repeat_region
/note="MLT1A1 repeat: matches 318. .365 of consensus"
repeat_region
/note="LNR7 repeat: matches 1. .450 of consensus"
repeat_region
/note="MLT1A1 repeat: matches 1. .318 of consensus"
repeat_region
/note="MIR repeat: matches 56. .142 of consensus"
repeat_region
/note="MIR repeat: matches 95. .146 of consensus"
repeat_region
/note="MIR repeat: matches 68. .212 of consensus"
repeat_region
/note="MIR repeat: matches 1. .711 of consensus"
repeat_region
/note="MER50 repeat: matches 13. .262 of consensus"
repeat_region
/note="MIR repeat: matches 13. .262 of consensus"
repeat_region
/note="Alusx repeat: matches 1. .304 of consensus"
repeat_region
/note="L2 repeat: matches 2506. .2695 of consensus"
repeat_region
10929. .11005

QY 1073 ACAGGCGCCTGAAGCTCAGAACCTCTAGACCCAGCTCCCTGAGATCACCA 1132
|||||
Db 1021 ACAGGCGCCTGAAGCTCAGAACCTCTAGACCCAGCTCCCTGAGATCACCA 1080
|||||
QY 1133 TCACCGAGGACCTGCTGGACACAGTGTAGCACTCCCTGCCCTGACCCAAATACCT 1192
|||||
Db 1081 TCACCGAGGACCTGCTGGACACAGTGTAGCACTCCCTGCCCTGACCCAAATACCT 1140
|||||
QY 1193 GCCAGCATGCCCGCCGCTGCTGCCCTGCTGGAGGTCCTCAACTGCCCTGGTCAATG 1252
|||||
Db 1141 GCCAGCATGCCCGCCGCTGCTGCCCTGCTGGAGGTCCTCAACTGCCCTGGTCAATG 1200
|||||
QY 1253 GCTCCCTCCACATCAGCAGAGCTGGTGCCTCATGATCAGGGGTCCCTGGCGCGGGC 1312
|||||
Db 1201 GCTCCCTCCACATCAGCAGAGCTGGTGCCTCATGATCAGGGGTCCCTGGCGCGGGC 1260
|||||
QY 1313 CTTGTGCTGCTGCTCCAGTGTCTCAAGATTGGGAGCAAGAAAGTCTCTACTGCT 1372
|||||
Db 1261 CTTGTGCTGCTGCTCCAGTGTCTCAAGATTGGGAGCAAGAAAGTCTCTACTGCT 1320
|||||
QY 1373 CTGAGCCACATCTCTCAAAAGGCCAACAAAAACAGAACTTCAAGTGGAGGTGATGG 1432
|||||
Db 1321 CTGAGCCACATCTCTCAAAAGGCCAACAAAAACAGAACTTCAAGTGGAGGTGATGG 1380
|||||
QY 1433 TACCAGTTAAAGCCAGTGGCAGCTCCCACTGCGGGCAGGCGAGACTGGAGTCTGTCTCA 1492
|||||
Db 1381 TACCAGTTAAAGCCAGTGGCAGCTCCCACTGCGGGCAGGCGAGACTGGAGTCTGTCTCA 1440
|||||
QY 1493 TTATGTATTCCTGGAGCCTGTGCTATGCTCACTGCAATTTTCAGGAGACATATTCAACT 1552
|||||
Db 1441 TTATGTATTCCTGGAGCCTGTGCTATGCTCACTGCAATTTTCAGGAGACATATTCAACT 1500
|||||
QY 1553 CCTGTGCTCTTCAAACTGGTGTCTATCCGGCAGAGGAGGAAGCAGACGAATGGAG 1612
|||||
Db 1501 CCTGTGCTCTTCAAACTGGTGTCTATCCGGCAGAGGAGGAAGCAGACGAATGGAG 1560
|||||
QY 1613 CAGGCTTTAGCAGAGCAGTTCCTGCGCAGAGCAGCCCTGCTGCCAGCGGGCAGCC 1672
|||||
Db 1561 CAGGCTTTAGCAGAGCAGTTCCTGCGCAGAGCAGCCCTGCTGCCAGCGGGCAGCC 1620
|||||
QY 1673 CCTCATAGGAGGCCAGGAGGAGCCCAAGGCGTAGAGCCCTGTTGAAGCTGTGAGCA 1732
|||||
Db 1621 CCTCATAGGAGGCCAGGAGGAGCCCAAGGCGTAGAGCCCTGTTGAAGCTGTGAGCA 1680
|||||
QY 1733 GGAGAACGGTGGCCACCAGCTTCAGAGTCTCCTGACCTGCTGCTATGCCCCAC 1790
|||||
Db 1681 GGAGAACGGTGGCCACCAGCTTCAGAGTCTCCTGACCTGCTGCTATGCCCCAC 1738
|||||

RESULT 6
LOCUS HS272L16 157875 bp DNA linear PRI 23-NOV-1999
DEFINITION Human DNA sequence from clone 272L16 on chromosome 1q32.1-32.3.
Contains the 3' end of the LAMB3 gene for Laminin, Beta 3 (Nicein, Kalinin, BM600) and a novel Rat Ca2+/Calmodulin dependent Protein Kinase LIKE gene. Contains ESTs, STSS, GSSs, genomic marker D1S491 and a ca repeat polymorphism, complete sequence.
ACCESSION AL023754
VERSION AL023754.1 GI:4007152
KEYWORDS HTG; BM600; ca repeat polymorphism; Ca2+/Calmodulin dependent Protein Kinase; D1S491; Kalinin; LAMB3; Laminin Beta 3; Nicein.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 157875)
Direct Submission
Submitted (27-NOV-1998) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
JOURNAL requests: clonerequest@sanger.ac.uk
On Dec 12, 1998 this sequence version replaced gi:3873472.

```
/note="L2 repeat: matches 2671. .2748 of consensus"
12092. .12129
repeat_region
/note="19 copies 2 mer ca 84% conserved"
12095. .12130
repeat_region
/note="9 copies 4 mer acac 86% conserved"
12722. .12855
repeat_region
/note="MER91A repeat: matches 20. .186 of consensus"
13108. .13221
repeat_region
/note="MIR repeat: matches 80. .191 of consensus"
complement(13503. .13979)
/note="match: GSSs AQ075457 AQ076848"
14235. .14279
repeat_region
/note="MIR repeat: matches 60. .105 of consensus"
14327. .14442
repeat_region
/note="MIR repeat: matches 112. .249 of consensus"
14495. .15046
repeat_region
/note="L2 repeat: matches 2127. .2750 of consensus"
16671. .16859
repeat_region
/note="MER5A repeat: matches 2. .189 of consensus"
17103. .17476
repeat_region
/note="L2 repeat: matches 17. .422 of consensus"
17679. .17938
repeat_region
/note="MIR repeat: matches 7. .262 of consensus"
17891. .17953
repeat_region
/note="L2 repeat: matches 2694. .2749 of consensus"
18650. .18758
repeat_region
/note="L2 repeat: matches 2641. .2750 of consensus"
19246. .19406
repeat_region
/note="MER5A repeat: matches 4. .189 of consensus"
19460. .19871
repeat_region
/note="L2 repeat: matches 2309. .2748 of consensus"
19916. .19987
repeat_region
/note="MIR repeat: matches 63. .140 of consensus"
20200. .20228
repeat_region
/note="MER34 repeat: matches 510. .538 of consensus"
20233. .20365
repeat_region
/note="MER34 repeat: matches 413. .543 of consensus"
20338. .20388
repeat_region
/note="LNR29 repeat: matches 454. .503 of consensus"
20459. .20809
repeat_region
/note="MER47A repeat: matches 2. .366 of consensus"
20976. .21143
repeat_region
/note="MER34 repeat: matches 6. .172 of consensus"
21834. .22278
repeat_region
/note="LNR1C repeat: matches 1. .466 of consensus"
22477. .22528
repeat_region
/note="26 copies 2 mer ag 79% conserved"
22826. .22937
repeat_region
/note="MER81 repeat: matches 1. .112 of consensus"
24897. .25034
repeat_region
/note="MER5A repeat: matches 30. .171 of consensus"
25084. .25193
repeat_region
/note="MIR repeat: matches 26. .145 of consensus"
25207. .25435
repeat_region
/note="MIR repeat: matches 13. .259 of consensus"
25919. .26363
repeat_region
/note="LNR1F repeat: matches 68. .541 of consensus"
26364. .26813
repeat_region
/note="MER42 repeat: matches 1. .503 of consensus"
26814. .26885
repeat_region
/note="LNR1F repeat: matches 1. .68 of consensus"
26931. .27357
repeat_region
/note="L2 repeat: matches 2292. .2723 of consensus"
27779. .27820
repeat_region
/note="21 copies 2 mer ca 100% conserved"
27779. .27818
repeat_region
/note="10 copies 4 mer caca 100% conserved"
29036. .29195
repeat_region
/note="HAL1 repeat: matches 4. .163 of consensus"
29263. .29500
repeat_region
/note="HAL1 repeat: matches 141. .355 of consensus"
29501. .29820
repeat_region
/note="LNR1A repeat: matches 1. .365 of consensus"
```

```
29821. .30076
/note="HAL1 repeat: matches 355. .618 of consensus"
30095. .30228
repeat_region
/note="FLAM_A repeat: matches 7. .142 of consensus"
30246. .30440
repeat_region
/note="HAL1 repeat: matches 683. .874 of consensus"
30539. .30788
repeat_region
/note="Charliela repeat: matches 1189. .1455 of consensus"
30789. .31099
repeat_region
/note="AlusJ repeat: matches 1. .306 of consensus"
31100. .31167
repeat_region
/note="Charliela repeat: matches 1125. .1189 of consensus"
31170. .31914
repeat_region
/note="L1M4 repeat: matches 5448. .6190 of consensus"
31923. .32141
repeat_region
/note="L1M2 repeat: matches 4317. .4530 of consensus"
32142. .32551
repeat_region
/note="MSTA repeat: matches 5. .426 of consensus"
32552. .33436
repeat_region
/note="L1M2 repeat: matches 3451. .4317 of consensus"
33437. .33756
repeat_region
/note="AluY repeat: matches 1. .309 of consensus"
33757. .36041
repeat_region
/note="L1M2 repeat: matches 761. .3451 of consensus"
37817. .38850
repeat_region
/note="Charliela repeat: matches 1. .1142 of consensus"
39293. .39437
repeat_region
/note="MIR repeat: matches 1. .140 of consensus"
complement(39554. .39839)
/note="match: GSS AQ070531"
41165. .41445
repeat_region
/note="L2 repeat: matches 2180. .2489 of consensus"
41615. .41810
repeat_region

Query Match          36.9%   Score 903;   DB 9;   Length 157875;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 903;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY 1535 CAGGAGACATATCAACTCCTCTCTTCCAAACCTGGTGTCTATCCGGCAGAGGAGG 1594
      |||||
DB 147496 CAGGAGACATATCAACTCCTCTCTTCCAAACCTGGTGTCTATCCGGCAGAGGAGG 147555

QY 1595 AAGCAGAGCAAGTGGAGCAGGCTTAGCAGGACAGTTCTGGCCAGAGACACAGCCT 1654
      |||||
DB 147556 AAGCAGAGCAAGTGGAGCAGGCTTAGCAGGACAGTTCTGGCCAGAGACACAGCCT 147615

QY 1655 GCTGCCAGCGGGCAGCGCCCTCATAGGAGGCCAGGAGGCCAGGCCAGGCGTAGAAGCC 1714
      |||||
DB 147616 GCTGCCAGCGGGCAGCGCCCTCATAGGAGGCCAGGAGGCCAGGCCAGGCGTAGAAGCC 147675

QY 1715 TTGTTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGC 1774
      |||||
DB 147676 TTGTTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCTCCCTGACCTGC 147735

QY 1775 CTGCTCTATGCCACACACCCCTACGTGCGTGGCTGTGTGCAGTGTACGTAGATAGCTCTC 1834
      |||||
DB 147736 CTGCTCTATGCCACACACCCCTACGTGCGTGGCTGTGTGCAGTGTACGTAGATAGCTCTC 147795

QY 1835 GCCTGGGTCTGTGCTGTTTGTCTGTAAGGCTTAATGGGCTGGCCAGGCTGTGTACACCTT 1894
      |||||
DB 147796 GCCTGGGTCTGTGCTGTTTGTCTGTAAGGCTTAATGGGCTGGCCAGGCTGTGTACACCTT 147855

QY 1895 CTCGAAGCAAGCCATATGGAGCATCTACCCAGACTCCCACTTCGCACACTCAGCTCC 1954
      |||||
DB 147856 CTCGAAGCAAGCCATATGGAGCATCTACCCAGACTCCCACTTCGCACACTCAGCTCC 147915

QY 1955 ACCTCTCAAGCTCCAACTCTTGGCCAGATTGGGCTCATTAATGTCTGCTTCCCTGCCCAT 2014
      |||||
DB 147916 ACCTCTCAAGCTCCAACTCTTGGCCAGATTGGGCTCATTAATGTCTGCTTCCCTGCCCAT 147975

QY 2015 CTGCATGAATGACAGGACGCTCCCATGGTGTGCTGTGAGCTCTTCAAGTTCTTAAT 2074
      |||||
DB 147976 CTGCATGAATGACAGGACGCTCCCATGGTGTGCTGTGAGCTCTTCAAGTTCTTAAT 148035
```

QY 2075 CCTTAAGTCCAGATTAGTCCCAAGTGGCTGAGACCCAGCCAGCACATCTTGGCCCT 2134
|||||
Db 148036 CCTTAAGTCCAGATTAGTCCCAAGTGGCTGAGACCCAGCCAGCACATCTTGGCCCT 148095
QY 2135 TCTCCCTGCTCAATCTAAAGCAGTGCACACCCCTCCAAAGTGAATAGAAAGATTC 2194
|||||
Db 148096 TCTCCCTGCTCAATCTAAAGCAGTGCACACCCCTCCAAAGTGAATAGAAAGATTC 148155
QY 2195 ATGAGTAAGGGTGAAGGAATCTTATCTGCGGCACATGCTCCGTCGACACACCCAA 2254
|||||
Db 148156 ATGAGTAAGGGTGAAGGAATCTTATCTGCGGCACATGCTCCGTCGACACACCCAA 148215
QY 2255 TGGAGTAACTTGAAGTGAATCTTATCTGCGGCACATGCTCCGTCGACACACCCCT 2314
|||||
Db 148216 TGGAGTAACTTGAAGTGAATCTTATCTGCGGCACATGCTCCGTCGACACACCCCT 148275
QY 2315 GTTCCCTTTCTCTCTCTGAAAGTCCAGCACACCATCTTCTGCTCCCGAGTTCTCTCG 2374
|||||
Db 148276 GTTCCCTTTCTCTCTGAAAGTCCAGCACACCATCTTCTGCTCCCGAGTTCTCTCG 148335
QY 2375 CCTCCACCCCTCCAGTTCATGCTAGTGTGCTGTTAATAAATGACATATTTTCT 2434
|||||
Db 148336 CCTCCACCCCTCCAGTTCATGCTAGTGTGCTGTTAATAAATGACATATTTTCT 148395
QY 2435 CTA 2437
|||
Db 148396 CTA 148398

RESULT 7
HUMD1S70R/c
LOCUS HUMD1S70R 1249 bp DNA linear PRI 17-MAR-1994
DEFINITION Human dinucleotide repeat polymorphism at the D1S70 locus.
ACCESSION L26254
VERSION L26254.1 GI:460666
KEYWORDS dinucleotide repeat polymorphism.
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1249)
AUTHORS Weston,M.D. and Kimberling,W.J.
TITLE A new polymorphism for the D1S70 locus
JOURNAL Unpublished (1993)
FEATURES
source
1..1249
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1q41"
748..771
/standard_name="dinucleotide repeat polymorphism"
/note="putative"
/rpt_type=tandem
repeat_region
BASE COUNT 310 a 229 c 391 g 319 t
ORIGIN
Query Match 3.2%; Score 78; DB 9; Length 1249;
Best Local Similarity 100.0%; Pred. No. 7.7e-33;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 628 GCTCCAGAGTGTCTGGCCAGAAACCTACAGCAAGGCTGTGGATTGCTGTGTCATCGGC 687
Db 987 GCTCCAGAGTGTCTGGCCAGAAACCTACAGCAAGGCTGTGGATTGCTGTGTCATCGGC 928
QY 688 GTCATCACTACATATG 705
|||||
Db 927 GTCATCACTACATATG 910
RESULT 8
BC021840
LOCUS BC021840 2416 bp mRNA linear ROD 07-AUG-2002

DEFINITION Mus musculus, clone MGC:30513 IMAGE:4502479, mRNA, complete cds.
ACCESSION BC021840
VERSION BC021840.1 GI:18256866
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2416)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 41 Row: i Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: GenomesScan gene
prediction.
FEATURES
source
Location/Qualifiers
1..2416
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:30513 IMAGE:4502479"
/tissue.type="Eye, retina, mouse strain C57BL/6"
/clone_lib="NIH_MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
87..1520
/codon_start=1
/product="Unknown (protein for MGC:30513)"
/protein_id="AAH21840.1"
/db_xref="GI:18256867"
/translation="MGRKEEDCSWKKQTNIKRTFTFMEVLGSGAFSEFLVKQRV
TGKFLAIKSPAFRODSLEIAVLKRIKHENIVTLEDIYEYTHYLVLMQLVSG
GELFDRIILRGVYTEKDALVQQVLSAVKYLHENGIVHRDILPENLYLTPENSKI
MITDFGLSKMEQNGVMSTACGTPGVAPVLAOKPYSKAVDCWSTGVITYILLCGYPP
FYEETESKLFEKIKFEGYEFESPFWDJISAKDFICHLEKDPNERVYTCFKALHPW
IDGNTALHRDIYPSVLSIQIKNFASKWRQAFNAAVYVHHMKLHMNLHSPSVQREVE
NRPPVPAPEVRPSDHSITEAPILDPPLPALTRLPCSHSSRSPSPAPSGGRSLNC
LVNGSLRISSLSLVPMQGPLATPGCGCCSCLNTGNKGKSYCSEPTTLFRKANKQNF
KSEVMVPYKAGSTHCRGGQTGCVLM"
BASE COUNT 572 a 694 c 602 g 548 t
ORIGIN
Query Match 2.9%; Score 71; DB 10; Length 2416;
Best Local Similarity 100.0%; Pred. No. 9.2e-29;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 304 AACATTGTGACCTGGAGGACATCTATGAGACACCCACCTACTACCTGTCATGCAG 363
|||||
Db 321 AACATTGTGACCTGGAGGACATCTATGAGACACCCACCTACTACCTGTCATGCAG 380
QY 364 CTGTTTCTCTGG 374
|||||
Db 381 CTGTTTCTCTGG 391


```
RESULT 9
AF428262
LOCUS
DEFINITION AF428262 2427 bp mRNA linear ROD 07-NOV-2001
Mus musculus calcium/calmodulin-dependent protein kinase I gamma
(Camk1g) mRNA, complete cds.
ACCESSION AF428262
VERSION
KEYWORDS
SOURCE AF428262.1 GI:16755793
Mus musculus.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2427)
AUTHORS Bjork.B.C., Watanabe.Y., Murray.J.C. and Schutte.B.C.
TITLE Characterization of the human ortholog of rat Cam Kinase I gamma
(CamK1g) at lq32-q41
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 2427)
AUTHORS Bjork.B.C., Watanabe.Y., Murray.J.C. and Schutte.B.C.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Genetics, University of Iowa, 140 EMRB,
Iowa City, IA 52242, USA
FEATURES
source
Location/Qualifiers
1..2427
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="1"
/map="1H, 104.00 cM"
/clone="IMAGE:4502479"
/tissue_type="retina"
/note="corresponds to EST BG293660"
1..2427
/gene="Camk1g"
86..1519
/gene="Camk1g"
/codon_start=1
/product="calcium/calmodulin-dependent protein kinase I
gamma"
/protein_id="AAL28101.1"
/db_xref="GI:16755794"
/translation="MGRKEEDCSWKKQTNIKRTIFMEVLGSGAFSEVLVKQPV
TKGLFALCKIKSPAFRDSLENIALVKRIKHENIVTLEDIYESTHYLYVMQVLSG
GELFDRIERGVYTKDASLYIQVLSAVKYLHENGIVHRDLKPENLLYITPENSKI
MITDGLSKMQNGVMSTACGTPGYVAPVLAQPKYSAVDWNSIGVITYILICGYPP
FYETESKLFEKIKEGYEFESFWDIDISEAKDFICHLLEKDPNERYTCEKALRHPW
IDGNTALHRDIYPSVLSIQKNEAKSKWROAFNAAVVHHMKLHMLHSPSVROEVE
NRPVSPAPEVSRDSDSHSITEAPILDPTPLPALTRLPCSHSRSPSAPSGGRSLNC
LYNGSLRTSSLSVPMQGGFLATGPGCCSCINIGNKGSYCISEPTLFRANKQNF
KSEVMVPVKAGGSTHCRGGGTGVCVM"
BASE COUNT 584 a 694 c 601 g 547 t 1 others
ORIGIN
Query Match 2.9%; Score 71; DB 10; Length 2427;
Best Local Similarity 100.0%; Pred. No. 9.2e-29;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 304 AACATTGTGACCTGGAGGACATCTATGAGACACACCCACTACTACTGTCATGCAG 363
|||||
Db 320 AACATTGTGACCTGGAGGACATCTATGAGACACACCCACTACTACTGTCATGCAG 379
|||||
QY 364 CTGTTCTCTGG 374
|||||
Db 380 CTGTTCTCTGG 390
|||||
RESULT 10
AL365314
LOCUS
DEFINITION AL365314 182054 bp DNA linear ROD 29-JUN-2002
Mouse DNA sequence from clone RP23-106N23 on chromosome 1, complete
sequence.
ACCESSION AL365314
VERSION
KEYWORDS
SOURCE
```

```
AL365314.16 GI:14422218
HTG.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 182054)
AUTHORS Kay.M.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk
On Jun 13, 2001 this sequence version replaced gi:13396595.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-106N23 is
from the RPCI-23 Mouse PAC library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
-----
Genome Center
Center: UK Medical Research Council
Web site: http://mrcseq.har.mrc.ac.uk
Contact: mouseq@har.mrc.ac.uk
-----
FEATURES
source
Location/Qualifiers
1..182054
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="1"
/clone="RP23-106N23"
/clone_lib="RPCI-23"
BASE COUNT 49155 a 44402 c 42300 g 46197 t
ORIGIN
Query Match 2.5%; Score 62; DB 10; Length 182054;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 304 AACATTGTGACCTGGAGGACATCTATGAGACACACCCACTACTACTGTCATGCAG 363
|||||
Db 84769 AACATTGTGACCTGGAGGACATCTATGAGACACACCCACTACTACTGTCATGCAG 84828
|||||
QY 364 CT 365
||
Db 84829 CT 84830
|||||
RESULT 11
AC022675/c
LOCUS
DEFINITION AC022675 184754 bp DNA linear HTG 09-MAY-2001
Mus musculus clone RP23-59N15, WORKING DRAFT SEQUENCE, 13 unordered
pieces.
ACCESSION AC022675
VERSION AC022675.3 GI:13940729
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus.
```


ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 184754)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-59N15

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

2 (bases 1 to 184754)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArleiano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferrelita,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneas,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

TITLE

JOURNAL

COMMENT

Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 4, 2001 this sequence version replaced gi:7657723.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5039
Center clone name: 59_N15

----- Summary Statistics
Sequencing vector: M13; M77815; 22% of reads
Sequencing vector: Plasmid; n/a; 78% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 179592 bases at least Q40
Consensus quality: 181970 bases at least Q30
Consensus quality: 182576 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 183554; sum-of-contigs
Quality coverage: 12.9 in Q20 bases; agarose-fp
Quality coverage: 12.8 in Q20.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. the true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 4354: contig of 4354 bp in length
* 4355 4454: gap of 100 bp
* 4455 5467: contig of 1013 bp in length
* 5468 5567: gap of 100 bp
* 5568 6717: contig of 1150 bp in length
* 6718 6817: gap of 100 bp
* 6818 8112: contig of 1295 bp in length
* 8113 8212: gap of 100 bp
* 8213 9269: contig of 1057 bp in length
* 9270 9369: gap of 100 bp
* 9370 11162: contig of 1793 bp in length
* 11163 11262: gap of 100 bp
* 11263 12591: contig of 1319 bp in length

* 12582 12681: gap of 100 bp
* 12682 13892: contig of 1211 bp in length
* 13893 13992: gap of 100 bp
* 13993 23297: contig of 9305 bp in length
* 23298 23397: gap of 100 bp
* 23398 38920: contig of 15523 bp in length
* 38921 39020: gap of 100 bp
* 39021 57984: contig of 18964 bp in length
* 57985 58084: gap of 100 bp
* 58085 141207: contig of 83123 bp in length
* 141208 141307: gap of 100 bp
* 141308 184754: contig of 43447 bp in length.

FEATURES

source

1. .184754
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="RP23-59N15"
/clone_lib="RP23-59N15" female Mouse BAC"

misc_feature

1. .4354
/note="assembly_fragment"
clone_end:SP6
vector_side:left"

misc_feature

4455. .5467
/note="assembly_fragment"

misc_feature

5568. .6717
/note="assembly_fragment"

misc_feature

6818. .8112
/note="assembly_fragment"

misc_feature

8213. .9269
/note="assembly_fragment"

misc_feature

9370. .11162
/note="assembly_fragment"

misc_feature

11263. .12581
/note="assembly_fragment"

misc_feature

12682. .13892
/note="assembly_fragment"

misc_feature

13993. .23297
/note="assembly_fragment"

misc_feature

23398. .38920
/note="assembly_fragment"

misc_feature

39021. .57984
/note="assembly_fragment"

misc_feature

58085. .141207
/note="assembly_fragment"

misc_feature

141308. .184754
/note="assembly_fragment"

misc_feature

clone_end:T7
vector_side:right"

BASE COUNT 47717 a 42686 c 43285 g 49746 t 1320 others

ORIGIN

Query Match 2.5%; Score 62; DB 2; Length 184754;
Best Local Similarity 100.0%; Pred.No.1.7e-23; Indels 0; Gaps 0;
Matches 62; Conservative 0; Mismatches 0;
QY 304 AACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCACTACTACCTGTCATGCAG 363
|||||
Db 140281 AACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCACTACTACCTGTCATGCAG 140222
QY 364 CT 365
Db 140221 CT 140220
RESULT 12
D86557 D86557 1013 bp mRNA linear ROD 07-FEB-1999
LOCUS Rattus norvegicus mRNA for Protein Kinase, partial cds.
DEFINITION D86557
ACCESSION D86557
VERSION D86557.1 GI:2077933
KEYWORDS Protein Kinase.
Rattus norvegicus embryo (E18) brain cDNA to mRNA, clone_lib:s.
SOURCE Nakanishi clone:N5.

```
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS      Yokokura,H., Terada,O., Naito,Y. and Hidaka,H.
TITLE        Isolation and comparison of rat cDNAs encoding
              Ca2+/calmodulin-dependent protein kinase I isoforms
JOURNAL      Biochim. Biophys. Acta 1338 (1), 8-12 (1997)
MEDLINE      97228532
REFERENCE
AUTHORS      Yokokura,H.
TITLE        Direct Submission
JOURNAL      Submitted (15-JUL-1996) Hisayuki Yokokura, Nagoya University School
              of Medicine, Department of Pharmacology; Tsurumai 65, Showa-ku,
              Nagoya, Aichi 466, Japan (Tel:052-744-2075, Fax:052-744-2083)
FEATURES
source
1. .1013
   /organism="Rattus norvegicus"
   /db_xref="taxon:10116"
   /clone="N5"
   /cissue_type="brain"
   /clone_lib="S. Nakanishi"
   /dev_stage="embryo (E18)"
   <1. .86
87. >1013
   /note="similar to calcium/calmodulin-dependent protein
       kinase I mRNA (L24907, L26288)"
   /codon_start=1
   /product="Protein Kinase"
   /protein_id="BAAL19880.1"
   /db_xref="GI:2077934"
   /translation="MGRKEEDCSWKKQTNIKRTFIEVLGSGAFSEVLVKQRV
       TGLFALCIKSPAFRDSSENEIIVLKRKHENIVTLEDIYESTHYLYVMQLVSG
       GELFRILRGVYTPKDSLVISAVKYLHENGIVHRDLKPENLLYLPPENSKI
       MTFDEGLSKMEONGYMGTCGTGPGVAPVLAQRPYKAVDCWIGVITYLLCGYPP
       FYEETSLKFIKIKGYEFESPFWDIDISEAKDFICHLEKDPNERYTCERKALHPW
       IDGNALHRDIVPSVSLQIQNFAKSKWRQAFN"
BASE COUNT   276 a 245 c 262 g 230 t
ORIGIN
Query Match      2.4%; Score 59; DB 10; Length 1013;
Best Local Similarity 100.0%; Pred. No. 8.6e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 778 GAGTTTGACCTCCATCTGGGATGACATTTCTGAGTCAGCCAGACGATTTATTGGCCA 836
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 795 GAGTTTGAGTCTCCATCTGGGATGACATTTCTGAGTCAGCCAGACGATTTATTGGCCA 853
RESULT 13
AC126166
LOCUS         AC126166
DEFINITION   Rattus norvegicus clone CH230-ID9, *** SEQUENCING IN PROGRESS ***,
              76 unordered pieces.
ACCESSION    AC126166
VERSION      AC126166.2 GI:21700365
KEYWORDS     HTG; HTGS; PHASE1.
SOURCE       Norway rat.
ORGANISM     Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS      Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
              Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,
              Barbara,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
              Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
              Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
              Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
              Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
              Cleveland,C.D., Cox,C., Coyte,M.D., Dathorne,S.R., David,R.,
              Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
              Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
              Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,
              Barbara,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
              Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
              Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
              Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
              Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
              Cleveland,C.D., Cox,C., Coyte,M.D., Dathorne,S.R., David,R.,
              Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
              Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
              Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
              Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
              Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
              Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
              Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
              Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
              Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
              Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
              Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S.,
              Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
              Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
              Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
              Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
              Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
              Massy,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
              Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
              Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
              Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,
              Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
              Peters,L., Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y.,
              Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
              Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
              Sodergren,E., Sonaik,A., Tabor,P., Tamerisa,A., Stanley,H., Stone,H.,
              Sutton,A., Svatek,A., Taylor,T., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
              Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
              Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
              Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K.,
              Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
              Weinstein,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 178272)
Worley,K.C.
Direct Submission
Submitted (04-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 178272)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 6, 2002 this sequence version replaced gi:21693754.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: TUCQ
Center clone name: CH230-ID9
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 113604 bases at least Q40
Consensus quality: 121900 bases at least Q30
Consensus quality: 127034 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
      (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
      consists of 76 contigs. The true order of the pieces
      is not known and their order in this sequence record is
      arbitrary. Gaps between the contigs are represented as
      runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
      as soon as it is available and the accession number will
      be preserved.
      1
      1071: contig of 1071 bp in length
      *
```

```
* 1072 1171: gap of unknown length
* 1172 2571: contig of 1400 bp in length
* 2572 2671: gap of unknown length
* 2672 4193: contig of 1522 bp in length
* 4194 4293: gap of unknown length
* 4294 5757: contig of 1464 bp in length
* 5758 5857: gap of unknown length
* 5858 6960: contig of 1103 bp in length
* 6961 7061: gap of unknown length
* 7062 8110: contig of 1050 bp in length
* 8111 9302: contig of 1092 bp in length
* 9303 9403: contig of 1092 bp in length
* 9404 10661: contig of 1259 bp in length
* 10662 10761: gap of unknown length
* 10762 12254: contig of 1493 bp in length
* 12255 12354: gap of unknown length
* 12355 13368: contig of 1014 bp in length
* 13369 13468: gap of unknown length
* 13469 14799: contig of 1331 bp in length
* 14800 14899: gap of unknown length
* 14900 15912: contig of 1013 bp in length
* 15913 16012: gap of unknown length
* 16013 17129: contig of 1117 bp in length
* 17130 17229: gap of unknown length
* 17230 18571: contig of 1342 bp in length
* 18572 18671: gap of unknown length
* 18672 20220: contig of 1449 bp in length
* 20221 20220: gap of unknown length
* 20221 21583: contig of 1363 bp in length
* 21584 21683: gap of unknown length
* 21684 23016: contig of 1333 bp in length
* 23017 23116: gap of unknown length
* 23117 24383: contig of 1167 bp in length
* 24384 24383: gap of unknown length
* 24384 25500: contig of 1117 bp in length
* 25501 25500: gap of unknown length
* 25501 25601: contig of 1709 bp in length
* 25601 27310: gap of unknown length
* 27310 27409: gap of unknown length
* 27409 29157: contig of 1748 bp in length
* 29158 29257: gap of unknown length
* 29258 31559: contig of 2302 bp in length
* 31560 31659: gap of unknown length
* 31660 32799: contig of 1140 bp in length
* 32800 32899: gap of unknown length
* 32900 35290: contig of 2391 bp in length
* 35291 35390: gap of unknown length
* 35391 36925: contig of 1535 bp in length
* 36926 37025: gap of unknown length
* 37026 38876: contig of 1851 bp in length
* 38877 38976: gap of unknown length
* 38977 40184: contig of 1208 bp in length
* 40185 40284: gap of unknown length
* 40285 41512: contig of 1228 bp in length
* 41513 41612: gap of unknown length
* 41613 42937: contig of 1325 bp in length
* 42938 43037: gap of unknown length
* 43038 43247: contig of 1210 bp in length
* 43248 44347: gap of unknown length
* 44348 46092: contig of 1745 bp in length
* 46093 46192: gap of unknown length
* 46193 47274: contig of 1082 bp in length
* 47275 47375: gap of unknown length
* 47375 48878: contig of 1504 bp in length
* 48879 48979: gap of unknown length
* 48979 50315: contig of 1337 bp in length
* 50316 50415: gap of unknown length
* 50416 52545: contig of 2130 bp in length
* 52546 52645: gap of unknown length
* 52646 54617: contig of 1972 bp in length
* 54618 54717: gap of unknown length
* 54718 55940: contig of 1223 bp in length
* 55941 56040: gap of unknown length

58002: contig of 1962 bp in length
58102: gap of unknown length
59810: contig of 1708 bp in length
59910: gap of unknown length
61918: contig of 2008 bp in length
62018: gap of unknown length
63896: contig of 1878 bp in length
63996: gap of unknown length
65703: contig of 1707 bp in length
65803: gap of unknown length
67650: contig of 1847 bp in length
67750: gap of unknown length
69467: contig of 1717 bp in length
69567: gap of unknown length
71839: contig of 2272 bp in length
71939: gap of unknown length
74071: contig of 2132 bp in length
74171: gap of unknown length
76853: contig of 2682 bp in length
76953: gap of unknown length
79190: contig of 2237 bp in length
79290: gap of unknown length
80794: contig of 1504 bp in length
80894: gap of unknown length
82989: contig of 2095 bp in length
83089: gap of unknown length
85392: contig of 2303 bp in length
85492: gap of unknown length
88314: contig of 2822 bp in length
88414: gap of unknown length
90655: contig of 2241 bp in length
90755: gap of unknown length
94227: contig of 3472 bp in length

Query Match 1.6%; Score 40; DB 2; Length 178272;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 778 GAGTTGAGTCTCCATTCTGGGATGACATTTCGTGAGTCAG 817
|||||
Db 91915 GAGTTGAGTCTCCATTCTGGGATGACATTTCGTGAGTCAG 91954

RESULT 14
AC120477/c
LOCUS AC120477 191584 bp DNA linear HTG 23-JUL-2002
DEFINITION Rattus norvegicus clone CH230-33N8, *** SEQUENCING IN PROGRESS ***,
71 unordered pieces.
ACCESSION AC120477.3 GI:21902854
VERSION AC120477
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 191584)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alzbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinb,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J.,
```

Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
 Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Lounseged,H.,
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mavhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okwuonu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rivas,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
 Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Tsmami,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstock,G. and Gibbs.R.
 Direct Submission
 Unpublished
 2 (bases 1 to 191584)
 Worley,K.C.
 Direct Submission
 Submitted (07-MAY-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 191584)
 Worley,K.C.
 Direct Submission
 Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 18, 2002 this sequence version replaced gi:20522179.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GXLL
 Center clone name: CH230-33N8
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 138862 bases at least Q40
 Consensus quality: 145657 bases at least Q30
 Consensus quality: 150087 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 71 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1006: contig of 1006 bp in length
 * 1007 1106: gap of unknown length
 * 1107 2111: contig of 1005 bp in length
 * 2112 3227: gap of unknown length
 * 3227 3328: contig of 1016 bp in length
 * 3328 4335: gap of unknown length
 * 4335 4435: gap of unknown length
 * 4435 5541: contig of 1006 bp in length
 * 5541 6587: gap of unknown length
 * 6587 8648: gap of unknown length
 * 8648 9951: contig of 1203 bp in length
 * 9951 10051: gap of unknown length
 * 10051 11081: contig of 1030 bp in length
 * 11081 12212: contig of 1031 bp in length
 * 12212 12313: gap of unknown length
 * 12313 13478: contig of 1166 bp in length
 * 13478 13578: gap of unknown length
 * 13578 14929: contig of 1351 bp in length
 * 14929 15029: gap of unknown length
 * 15029 16090: contig of 1061 bp in length
 * 16090 16190: gap of unknown length
 * 16190 17244: contig of 1054 bp in length
 * 17244 17344: gap of unknown length
 * 17344 18956: contig of 1612 bp in length
 * 18956 19056: gap of unknown length
 * 19056 20171: contig of 1115 bp in length
 * 20171 20271: gap of unknown length
 * 20271 21283: contig of 1012 bp in length
 * 21283 21383: gap of unknown length
 * 21383 23582: contig of 2199 bp in length
 * 23582 24984: contig of 1302 bp in length
 * 24984 25084: gap of unknown length
 * 25084 26255: contig of 1171 bp in length
 * 26255 26355: gap of unknown length
 * 26355 27622: contig of 1267 bp in length
 * 27622 27723: gap of unknown length
 * 27723 29025: contig of 1303 bp in length
 * 29025 29125: gap of unknown length
 * 29125 30966: contig of 1841 bp in length
 * 30966 31066: gap of unknown length
 * 31066 32073: contig of 1007 bp in length
 * 32073 32173: gap of unknown length
 * 32173 33601: contig of 1428 bp in length
 * 33601 33701: gap of unknown length
 * 33701 34879: contig of 1178 bp in length
 * 34879 34979: gap of unknown length
 * 34979 36112: contig of 1133 bp in length
 * 36112 36212: gap of unknown length
 * 36212 38536: contig of 2324 bp in length
 * 38536 38636: gap of unknown length
 * 38636 40088: contig of 1452 bp in length
 * 40088 40188: gap of unknown length
 * 40188 43004: contig of 2816 bp in length
 * 43004 43104: gap of unknown length
 * 43104 4472: contig of 1368 bp in length
 * 4472 4473: gap of unknown length
 * 4473 47031: contig of 2459 bp in length
 * 47031 47131: gap of unknown length
 * 47131 49092: contig of 1961 bp in length
 * 49092 49192: gap of unknown length
 * 49192 51804: contig of 2612 bp in length
 * 51804 51904: gap of unknown length
 * 51904 53859: contig of 1955 bp in length
 * 53859 53959: gap of unknown length
 * 53959 55351: contig of 1392 bp in length
 * 55351 55451: gap of unknown length
 * 55451 57797: contig of 2346 bp in length
 * 57797 60119: contig of 2222 bp in length
 * 60119 60219: gap of unknown length
 * 60219 61884: contig of 1665 bp in length
 * 61884 61984: gap of unknown length
 * 61984 65400: contig of 3416 bp in length
 * 65400 65500: gap of unknown length
 * 65500 67405: contig of 1905 bp in length

* 67406 67505: gap of unknown length
* 67506 68550: contig of 1145 bp in length
* 68551 68750: gap of unknown length
* 68751 70564: contig of 1814 bp in length
* 70565 70664: gap of unknown length
* 70665 72651: contig of 1987 bp in length
* 72652 72751: gap of unknown length
* 72752 75008: contig of 2257 bp in length
* 75009 75108: gap of unknown length
* 75109 77573: contig of 2465 bp in length
* 77574 77674: gap of unknown length
* 77675 79219: contig of 1546 bp in length
* 79220 79319: gap of unknown length
* 79320 81724: contig of 2405 bp in length
* 81725 81824: gap of unknown length
* 81825 84362: contig of 2538 bp in length
* 84363 84462: gap of unknown length
* 84463 86195: contig of 1733 bp in length
* 86196 86295: gap of unknown length
* 86296 89725: contig of 3430 bp in length
* 89726 89825: gap of unknown length
* 89826 93336: contig of 3511 bp in length
* 93337 93437: gap of unknown length
* 93437 96552: contig of 3116 bp in length
* 96553 96552: gap of unknown length
* 96553 99918: contig of 3266 bp in length

Query Match 1.6%; Score 40; DB 2: Length 191584;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 778 GAGTTTGAGTCTCCATTCGGATGACATTTCTGAGTCAG 817
|||||
Db 49589 GAGTTTGAGTCTCCATTCGGATGACATTTCTGAGTCAG 49550

RESULT 15

BC014825
LOCUS BC014825 1448 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, Similar to calcium/calmodulin-dependent protein kinase I, clone MGC:18933 IMAGE:3969343, mRNA, complete cds.
ACCESSION BC014825
VERSION BC014825.1 GI:15928725
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1448)
Strausberg, R.
Direct Submission
Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

REMARK COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 23 Row: j Column: 19

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES source

Location/Qualifiers

1..1448
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:18933 IMAGE:3969343"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Maml"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
120..1244
/codon_start=1
/product="Similar to calcium/calmodulin-dependent protein kinase I"
/protein_id="AAH14825.1"
/db_xref="GI:15928726"
/translation="MPGAVEGPRWQAEIRIDYDFRDLVLTGAFSEVILAEDEKRTQK LVAIKCIATKALEGKESMENEIATLHKIKHNIVALDDDIYESGGHLYLIMQLVSGGE LFDRIIVERKEGYETDASRLIFQVLDAAVYKLDGIVHRDLKPENLLYSLDDSDSKIMI SDFGLSKMEDPGSVLSTAGTGYVAPEVLAOKPYSKAVDCWSGVIAVILLCGYPPF YDENDAKLFEQILKAEYEDSPYWDIDSDAKDFIRHLMEKDPKRFCEQALQHPWI AGDTALDKNIHQSVSEQIKKFKAKSKWQAFNATAVVRHMKLQLGTSGQGGQGTGSH GELLTPTAGGPAAGCCCRCCVPEPSSELPAPPSSRAMD"

BASE COUNT 364 a 375 c 408 g 301 t
ORIGIN

Query Match 1.3%; Score 32; DB 10: Length 1448;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 CCCTACAGCAAGGCTGTGGATTGCTGCTCCAT 683
|||||

Db 699 CCCTACAGCAAGGCTGTGGATTGCTGCTCCAT 730
|||||

Search completed: March 15, 2003, 04:12:48
Job time : 8188 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2003, 04:12:52 ; Search time 2223 Seconds
(without alignments)
3467.860 Million cell updates/sec

Title: US-09-960-643-2
Perfect score: 2513
Sequence: 1 MGRKEEDDCSSWKKQTNR.....VKAGSSSHCRAGQTGVCLIM 476

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09960643/runat_07032003_090517_19885/app_query.fasta_1.647
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09960643 @CGN.1.1.1906 @runat_07032003_090517_19885 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPEXT=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_estl.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rtd.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1532	61.0	1126	13	BM547443
2	1298	51.7	905	14	BQ934044
3	1257	50.0	809	13	BI821474
4	1234.5	49.1	812	13	BI772626
5	1191.5	47.4	1129	14	BM807335
6	1158	46.1	1051	9	AL560091
7	1147	45.6	740	13	BI824483
8	1140.5	45.4	1068	14	BM921532
9	1134	45.1	742	13	BI818261
10	1122	44.6	731	13	BI753035
11	1103	43.9	818	12	BG293660
12	1101	43.8	824	12	BG715920
13	1013.5	40.3	918	14	BQ949629
c	985	39.2	555	10	AW251224
14	966.5	38.5	962	9	AL556476
15	938.5	37.3	620	14	BM944418
16	920.5	36.6	573	14	BQ866330
17	901.5	35.9	1068	14	BM919261
18	894.5	35.6	873	12	BG168668
19	893.5	35.6	554	13	BI834635
20	879.5	35.0	929	13	BI758466
21	871	34.1	871	10	BE569540
22	856.5	34.0	1052	14	BQ666920
23	855.5	33.8	656	13	BI667965
24	849	33.3	627	10	AW826802
25	836	33.0	973	13	BM051435
26	830.5	32.7	822	13	BI254545
27	820.5	32.4	822	13	BI545545
c	813	32.3	798	12	BG702279
28	812.5	32.3	798	12	BG702279
29	800	31.8	659	13	BI463345
30	797.5	31.7	607	9	AI595095
31	797.5	31.7	817	13	BI756289
32	797.5	31.7	725	12	BG302505
33	793	31.6	636	10	BE266955
34	793	31.6	941	9	AL578723
c	790	31.4	439	10	AW522224
35	785	31.2	863	13	BI757461
36	784.5	31.2	863	13	BI757461
37	780.5	31.1	646	14	BM728430
38	780.5	30.9	832	12	BF582563
39	777.5	30.7	748	12	BF143967
40	772.5	30.7	748	12	BG895406
41	766.5	30.5	556	12	BQ897490
42	766	30.5	929	14	BM903712
43	765	30.4	1061	14	BM903712
44	763.5	30.4	758	13	BI761267
45	759.5	30.2	640	12	BG307424

ALIGNMENTS

RESULT 1
BM547443
LOCUS
DEFINITION AGENCOURT_6507688 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5724450
5', mRNA sequence.
ACCESSION BM547443
VERSION BM547443.1 GI:18781213
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1126)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgraphs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12713 row: 1 column: 19
High quality sequence stop: 623.
Location/Qualifiers
1. .1126
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5724450"
/clone_lib="NIH_MGC_125"
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6; Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool of three ovaries, from females ranging in age from 38 to 49 yo. library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 2.1 kb, insert size range 1-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 036."
BASE COUNT 299 a 291 c 286 g 244 t 6 others
ORIGIN

FEATURES
source

Alignment Scores:

Pred. No.: 3,44e-146 Length: 1126
Score: 1532.00 Matches: 316
Percent Similarity: 90.70% Conservative: 6
Best Local Similarity: 89.01% Mismatches: 27
Query Match: 60.96% Indels: 7
DB: 13 Gaps: 2

US-09-960-643-2 (1-476) x BM547443 (1-1126)

Qy 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrTrpAsnIleArg 20
Db 69 ATGGTTCGAAGGAAGAGATGACTGCTGCTGGAAGAACAGACACCATCCGG 128
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 129 AAAACCTTCATTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGTGTTCCTGGTG 188
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 189 AAGCAAGACTGACTGGGAAGCTTTTGCTCTGAAGTGCATCAAGAGTCACTGCCTTC 248
Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysLysLysGluAsnIle 80
Db 249 CGGACAGACGCTGGGAATGAGATGCTGCTGTGGAAGATCAAGCATGAACAT 308
Qy 81 ValThrLeuGluAspIleTrpGluSerThrThrHisTrpTyrLeuValMetGlnLeuVal 100
Db 309 GTGACCTGGAGGACATCTATGAGAGCACCACCTACTACCTGCTGTCATGAGCTGT 368
Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValThrGluLysAspAla 120
Db 369 TCTGTGGGAGCTCTTTGACCGGATCCTTGGAGCGGGGTCTTACACAGAGAGGATGCC 428
Qy 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTrpLeuHisGluAsnGlyIleVal 140
Db 429 AGTCTGTGTATCCAGCAGGCTTGTGCGCAGTGAATACCTACATGAGATGTCATCGTC 488
Qy 141 HisArgAspLeuLysProGluAsnLeuLeuThrProGluGluAsnSerLysIle 160
Db 489 CACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCTTGAAGAACTCTAAGATC 548

Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 549 ATGATCACTGACTTTTGGTCTGTCCAAGATGGACAGAGATCATGTCCACTGCGTGT 608
Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 609 GGGACCCAGAGCTACGTGGCTCCAGAGTGTGCGCCAGAAACCTACAGCAAGCGTGTG 668
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220
Db 669 GATTGCTGCTCATCGGCTCATCACCTACATATTGCTCTGTGATACCCCATTTCTAT 728
Qy 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Db 729 CAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGTTTGAGTCT 788
Qy 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHis-LeuLeuGluLys 260
Db 789 CATTCTCGGATGACATTCTTGAGTCACCAAGGACTTTATTGCCACTTTGCTTGAGAA 848
Qy 260 sAspProAsnGluArgTyrThr-CysGluLysAlaLeu---SerHisProTrpIleAspG 279
Db 849 AGATCCGAACGAGCGGTACACCTGNTGAGAAAGGCTTGAGTCATCCCTGGGATTGACG 908
Qy 279 LysThrAla-LeuHisArg-AspIleTyrProSerValSerLeuGlnIleGlnLysAs 298
Db 909 GAAACACAGCCCTCCACCGGNGACATCTACCATCATGCTGAGCTTCAGATCCAGAGAA 968
Qy 298 nPheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValAlaHisHisMetAr 318
Db 969 CTGTGCTAANAACCAATGGAGGAGGAGCTTCAACCCANCAGCTGTGGTGGCCCCCTGAG 1028
Qy 318 gLysLeuHisMetAsnLeuHisSerProGlyValArgProGluVal---GluAsnArgPr 337
Db 1029 GAAAGTACC-CTGAACCTGGACAGCGGGAATTCCNNCCCCAAAGGGGGGAAACAGGCC 1087
Qy 337 ProGluThrGlnAlaSerGluThrSerArgProSer 349
Db 1088 GCCCTGAAATCCAGCCCTCAAAACCTTTGAAACCCAGT 1124

RESULT 2

LOCUS B0934044 905 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8802332 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6336727
ACCESSION B0934044
VERSION B0934044.1 GI:22349427
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 905)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgraphs-r@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13801 row: d column: 08
High quality sequence stop: 689.
Location/Qualifiers
1. .905
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6336727"

FEATURES
source

/clone_lib="NIH_MGC_130"
/lab_host="DH10B (phage-resistant)"
/note="Organ: otocysts; Vector: pCMV-SPORT6.1.cdb;
Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally.
Primer: Oligo dt. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is a
NIH_MGC Library."
BASE COUNT 238 a 217 c 240 g 210 t
ORIGIN

Alignment Scores:
Pred. No.: 2,42e-122 Length: 905
Score: 1298.00 Matches: 256
Percent Similarity: 98.12% Conservative: 5
Best Local Similarity: 96.24% Mismatches: 1
Query Match: 51.65% Indels: 4
DB: 14 Gaps: 0
US-09-960-643-2 (1-476) x BQ934044 (1-905)

QY 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrAsnLeuArg 20
|||||
Db 107 ATGGGGCGTAAGGAGGAGGAGGAGTCTGCTGCTGGAAGAAACAGACCAACATCAGG 166
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
|||||
Db 167 AAAACCTTCATCTTCATGAAGTGCCTGGGATCAGGAGCTTCTCAGAGGTGTCTCTGGTG 226
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
|||||
Db 227 AAGCAAAAGATGACTGGGAACTCTTTGCTGCTGAAATGATATCAAGAGTACCAGCCTTC 286
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
|||||
Db 287 CGGACAGACAGCCTAGAGATGAGATCGCTGTGTTGAAAGAGATCAAGCATGAGAACAT 346
QY 81 ValThrLeuGluAspIleTrpGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
|||||
Db 347 GTGACCTGGAGGACATCTATGAGAGCACCACCTACTACTCTGTCATGCGAGTTGTT 406
QY 101 SerGlyGlyCyluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
|||||
Db 407 TCTGGAGGTGAGCTTTTACCGGATCCTAGAGCGTGGTGTCTACACAGAAAGAGATGCC 466
QY 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
|||||
Db 467 AGCCTGTCATCAGCAGGCTGTGCTGCGGTGAATACCTTCATGAGATGGCATCGTC 526
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
|||||
Db 527 CACAGAGATCTAAAGCCTGAAACCTGCTGTACCTCACCCCTGAGGAGAACTCCCAAGATC 586
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
|||||
Db 587 ATGATCACTGACCTTGGTCTATCCAGATGGAGATGGAGTATGTCACACCTGTGT 646
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
|||||
Db 647 GGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCAGCAAGCCCTACAGTAAGGCTGTG 706
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuLysGlyTyrProPheThr 220
|||||
Db 707 GACTGCTGGTCCATTTGGTGTACATACATACACTGCTGTGGGTATATCCCTTTTCTAT 766
QY 221 GluGluThrGluSerLysLeu-PheGluLysIleLysGluGlyTyrTyrGlu-PheGlu 240
|||||
Db 767 GAAGAAACAGATCAAGCTTTTTCAGAGATCAAAAGAGGTACTACGAGGTTTGAGT 826
QY 240 erProPheTrpAspIleSerGluSer-AlaLysAspPheIleCysHis-LeuLeuGlu 259
|||||
Db 827 CTCGGTCTGGGATGACATTTCTGAGTCAGGCCAGGGATTTTATTTATGTCATCCTCTGGA 886
QY 259 uLysAspPro 262

|||||
Db 887 GAGGACCCC 896

RESULT 3
BI821474
LOCUS
DEFINITION
603038366F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179336 5',
mRNA sequence.
ACCESSION
BI821474
VERSION
BI821474.1 GI:15933024
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 809)
NIH-MGC http://mgc.nci.nih.gov/
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@rsr@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1447 row: c column: 17
High quality sequence stop: 808.
FEATURES
Location/Qualifiers
source
1..809
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5179336"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 212 a 192 c 218 g 187 t
ORIGIN
Alignment Scores:
Pred. No.: 3,24e-118 Length: 809
Score: 1257.00 Matches: 241
Percent Similarity: 99.59% Conservative: 1
Best Local Similarity: 99.18% Mismatches: 1
Query Match: 50.02% Indels: 0
DB: 13 Gaps: 0
US-09-960-643-2 (1-476) x BI821474 (1-809)

QY 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrAsnIleArg 20
|||||
Db 75 ATGGGTGCAAGGAGGAGATGATGCTGAGTCTGGAAGAAACAGACCAACATCCCG 134
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
|||||
Db 135 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAGTTTCTCTGGTG 194
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
|||||
Db 195 AAGCAAAAGACTGACTGGGAGGCTTTCGCTCTGAGTGCATCAAGAGTACACCTGCTTC 254
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80


```
Db 722 CCATTTCGGATGACATTCTGAGTCAGCCAGGAGCTTTATTGGCCACTTGTCTTGAGAAG 781
|||||
QY 261 AspProAsnGluArgTyrThrCysGluLys 270
|||||
Db 782 GATCCGACGAGCGGTACACCTGTGAGAAG 811
|||||

RESULT 5
BM807335
LOCUS BM807335 1129 bp mRNA linear EST 05-MAR-2002
DEFINITION AGENCOURT 6575096 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5732298
5' mRNA sequence.
ACCESSION BM807335
VERSION BM807335.1 GI:19124158
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 1129)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@ncl.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM2734 row: c column: 19
High quality sequence start: 11
High quality sequence stop: 651.
FEATURES
source
1. 1129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5732298"
/clone_lib="NIH_MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site: 1: EcoRV
(destroyed); Site: 2: NotI; RNA source male hippocampus,
age 27; library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 287 a 308 c 294 g 240 t
ORIGIN
Alignment Scores:
Pred. No.: 2.89e-111 Length: 1129
Score: 1191.50 Matches: 281
Percent Similarity: 72.86% Conservative: 9
Best Local Similarity: 70.60% Mismatches: 40
Query Match: 47.41% Indels: 70
DB: 14 Gaps: 7

US-09-960-643-2 (1-476) x BM807335 (1-1129)
QY 1 MetGlyArgLysGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
|||||
Db 105 ATGGGTCTGAAGGAGAGATGATCGATGCTCTGGGAAGAACAGACCAACATCCCG 164
|||||
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
|||||
Db 165 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGTTTCTCGGTG 224
|||||
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
|||||
```

```
Db 225 AAGCAAAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCAACCTGCCTTC 284
|||||
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
|||||
Db 285 CGGACAGACGCCCTGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAACATT 344
|||||
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
|||||
Db 345 GTGACCTGGAGGACATCTATGAGAGCACCACCACCTACTACCTGGTGCATGCAGCTGTGT 404
|||||
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
|||||
Db 405 TCTGGTGGGAGCTCTTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGATGCC 464
|||||
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
|||||
Db 465 AGTCTGGTGATCCAGCAGGCTTGTTCGCAGTGAAATACCTACATGAGATGCGATCGTC 524
|||||
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrThrProGluGluAsnSerLysIle 160
|||||
Db 525 CACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCTGAAGAGAACTCTAAGATC 584
|||||
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
|||||
Db 585 ATGATCACTGACTTTGGTCTGTCCAAGATGAACAGATGCGATCATGTCCTCCATGCTGT 644
|||||
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAla-Va 200
|||||
Db 645 GGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCCGAGAAACCTACAGCAAGGCTGT 704
|||||
QY 200 Lasp-CysTrp-SerIleGlyValIleThr-TyrIleLeuLeuCysGly---TyrProPr 218
|||||
Db 705 GGATTTGCTGGTCCATCGGGGGCATCACCTACCTATTGCTCTGTGATTAACCCCTCC 764
|||||
QY 218 oPheTyrGlu-GluThrGluSerLysLeuPheGluLysIleLysGluGlyTyr---TyrG 237
|||||
Db 765 GTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAAGGGCTAATGTA 824
|||||
QY 237 LuPheGluSerProPheTrp-AspAsp-IleSerGluSerAla-LysAspPhe-IleCys 255
|||||
Db 825 GTTTTGAGTCTCCATTCTGGGGATGAACATTCCTGAGTCAGCCCGAGGAGCTTTAATTGC 884
|||||
QY 256 HisLeu---LeuGluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHis 274
|||||
Db 885 CACCTTGCTTTGAAAAGGATCCCAACGAGGGGGACCCCGG----- 929
|||||
QY 275 ProTrpIleAspGlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGln 294
|||||
Db 930 -----GGGAACAGGCGCTTTGAG----- 947
|||||
QY 295 IleGlnLysAsnPheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValVal 314
|||||
Db 947 ----- 947
|||||
QY 315 HisIleMetArgLysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGlu 334
|||||
Db 948 -----CACTCCCGGGGACA-TGGACGGAAAAACACT 976
|||||
QY 335 AsnArgProPro-----GluThrGlnAlaSerGluThrSerArgProSerProGlu 352
|||||
Db 977 GGCCCTCCCGGGGGGCCACTCCCGCTTCATATAAGCCT-CCCCAAAATCCAGAG 1035
|||||
QY 353 IleThrIleThrGluAlaProValLeuAspHisSerValAlaLeuPro-AlaLeuThrG1 372
|||||
Db 1036 AAAACTTTGTCC-----ATACCACCAAGTGGCGACGCCAAGCGCTTTAAC 1083
|||||
QY 372 nLeuProCysGlnHisGlyArgArgProThrAla 383
|||||
Db 1084 GGCCCCAGGAGCTCGTGGCGCGGCCCTCTCT 1117
|||||
RESULT 6
AL560091
```

LOCUS AL560091 1051 bp mRNA linear EST 16-FEB-2001
 DEFINITION AL560091.LTI_FL011_BCI Homo sapiens cDNA clone CS0DG002YN16 5 prime
 , mRNA sequence.

ACCESSION AL560091
 VERSION AL560091.1 GI:12906218
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1051)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 1051

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DG002YN16"

/clone_lib="LTI_FL011_BCI"

/sex="male"

/tissue_type="B cells from Burkitt lymphoma"

/lab_host="DH10B"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-stranded cDNA was digested with Not I and cloned
 into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
 Library was constructed by Life Technologies. Contact :
 Feng Liang Life Technologies, a division of Invitrogen
 9800 Medical Center Drive Rockville, Maryland 20850, USA
 Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 274 a 270 c 287 g 205 t 15 others

ORIGIN

Alignment Scores:

Pred. No.: 7.02e-108 Length: 1051

Score: 1158.00 Matches: 226

Percent Similarity: 84.29% Conservative: 37

Best Local Similarity: 72.44% Mismatches: 47

Query Match: 46.08% Indels: 3

DB: 9 Gaps: 2

US-09-960-643-2 (1-476) x AL560091 (1-1051)

Qy 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20

Db 111 ATGCCCGGGAGACGGCGAGAGCAGCTCTCTGGAAAAGCAAGCTGAAGACATCAAG 170

Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40

Db 171 AAGATCTTCGAGTTCAAGAGACCCTCGAACCAGGCGCTTTCCGAAGTGGTTTAGCT 230

Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59

Db 231 GAAGAGAAGGCAACTGGCAAGCTCTTTGCTGTGAAGTGTATCCCTAAGAGGCGCTGAAG 290

Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79

Db 291 GCGAAGGAACACATAGAGATGAGATAGCCGCTCTGAGAAGATTAAGCATGAATAAT 350

Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99

Db 351 ATTTGTCCTCGGAAGACATTTATGAAGCCCAATACCTGACTTGGTCATGCAGCTG 410

Qy 100 ValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119

Db 411 GTGTCGGTGGAGAGCTGTTTGACCGGATAGTGGAAGAGGGGTTTATACAGAGAAGGAT 470

Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139

Db 471 GCCACACTCTGATCCGCCAAGTCTTGGACGCGGTGTACTATCCACAGAATGGCATC 530

Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159

Db 531 GTCCACAGAGAGCTCAAGCCGAAATCTCTGTACTACAGTCAAGATGAGAGATCCAAA 590

Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178

Db 591 ATAATGATCAGTGACTTGGATTGTCAAAATAGGAGCAAGAGATGTGATGTCCACT 650

Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198

Db 651 GCCTGTGGAACCTCCAGGCTATGTGCTCTGAAGTCCTCGCCAGAAACCTTACAGCAA 710

Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218

Db 711 GCGGTGTACTGCTGTGTCATCGGAGTGATGCTTGTCTTGGCGCTACCCGCT 770

Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238

Db 771 TTTTATGATGAATGACTCCAGCTCTTGAGCAGATCCTCAAGCGCAATATGAGTTT 830

Qy 239 GluSerProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258

Db 831 GACTCTCCCTACTGGATGACATCTCCGACTCTGCAAAAGACTTCTTCGCGAATGATG 889

Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278

Db 890 GAGAAGGCCCAATGAAGATACACGTGTGAGCAGGAGCTCGGCACCCATGATCGCT 949

Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298

Db 950 GGTGACACAGCCCTCAAAATAAACATCCAGGAGTCGTCAGCGCCAGATCCGGAATAAC 1009

Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAla 310

Db 1010 TTTGCCAAGAGCAATGGAGCAAGCATTTTATGCC 1045

RESULT 7

LOCUS BI824483 740 bp mRNA linear EST 04-OCT-2001

DEFINITION 603038855F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179957 5',

ACCESSION BI824483

VERSION BI824483

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 740)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Contact: cgapps-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11448 row: m column: 14

High quality sequence stop: 709.

Location/Qualifiers

1. 740

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5179957"

/clone_lib="NIH_MGC_115"

FEATURES

SOURCE

```
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pCMW-SPOK16; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT      196 a   181 c   196 g   167 t
ORIGIN

Alignment Scores:
Pred. No.:      5,37e-107      Length:      740
Score:          1147.00      Matches:      222
Percent Similarity: 99.55%      Conservative: 0
Best Local Similarity: 99.55%      Mismatches: 1
Query Match:      45.64%      Indels:      0
DB:              13      Gaps:      0

US-09-960-643-2 (1-476) x BI824483 (1-740)

Qy 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 70 ATGGGTCGAAGAAGAGAGATGACTGCTGCTGGAAGAAGAACACACCAACATCCGG 129
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 130 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAGTTTCTCGTGTG 189
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 190 AAGCAAGACTGACTGGGAGCTCTTTGCTCTGAAGTGCATCAAGAACTCAGCTCCCTTC 249
Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysLysIleLysHisGluAsnIle 80
Db 250 CGGACAGCAGCTGGAGATGAGATGCTGTGTTGAAAAGATCAAGCATCAAAACATT 309
Qy 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 310 GTGACCTGGAGGACATCTATGAGAGCACCACCCACTACTGCTGTCATGCGAGCTTGT 369
Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Db 370 TCTGGTGGGAGCTCTTGACCGGATCCTGGAGCGGGGTCTACACAGAGAAGATGCC 429
Qy 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 430 AGTCTGTGATCCAGCAGGCTTGTGCGGAGTGAATACTACATGAGAATGGCATCGTC 489
Qy 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 490 CACAGACATTAAAGCCCGAAACCTGCTTACCTTACCCCTGAAGAGAACTCTAAGATC 549
Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 550 ATGATCACTGACITTTGTCTGTCCAAGATGGAACAGATGGCATCATGTCCTCCCTGT 609
Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 610 GGGACCCAGGCTACGTGGCTCCAGAAGTGTCTGGCCCAAGAAACCCCTACACAGAGCTGTG 669
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLysCysGlyTyrProProPheTyr 220
Db 670 GATTGCTGTTCATCGCGCTCATCACTACATATTGCTCTGGGATACGCCCCCGTTCTAT 729
Qy 221 GluGluThr 223
Db 730 GAAGAAACG 738

RESULT 8
```

```
BM921532
LOCUS
DEFINITION
  AGENCOURT_6708041 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753010
  5', mRNA sequence.
ACCESSION
  BM921532
VERSION
  BM921532.1 GI:19371911
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 1068)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cygabs-r@mail.nih.gov
  Tissue Procurement: Life Technologies, Inc.
  CDNA Library Preparation: Life Technologies, Inc.
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: L1AM12788 row: b column: 19
  High quality sequence stop: 592.
  Location/Qualifiers
    source
      1..1068
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5753010"
        /clone_lib="NIH_MGC_115"
        /lab_host="DH10B"
        /note="Organ: pooled brain, lung, testis; Vector:
        pCMW-SPOK16; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
        source anonymous pool of 6 male brains, age range 23-27; 1
        male lung, age 27; and 1 male testis, age 69. Library is
        oligo-dT primed and directionally cloned (EcoRV site is
        destroyed upon cloning). Average insert size 1.8 kb,
        insert size range 1-3 kb. Library is normalized and
        enriched for full-length clones and was constructed by C.
        Gruber (Invitrogen). Research Genetics tracking code
        021. Note: this is a NIH_MGC Library."
BASE COUNT      279 a   281 c   260 g   248 t
ORIGIN

Alignment Scores:
Pred. No.:      4.49e-106      Length:      1068
Score:          1140.50      Matches:      229
Percent Similarity: 94.67%      Conservative: 2
Best Local Similarity: 93.85%      Mismatches: 10
Query Match:      45.38%      Indels:      3
DB:              14      Gaps:      1

US-09-960-643-2 (1-476) x BM921532 (1-1068)

Qy 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 83 ATGGGTCGAAGAAGAGATGACTGCTGCTGGAAGAAGAACACACCAACATCCGG 142
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 143 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAGTTTCTCGTGTG 202
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 203 AAGCAAGACTGACTGGGAGCTCTTTGCTCTGAAGTGCATCAAGAGTCAACCTGCTTC 262
Qy 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
Db 263 CGGACAGCAGCCTGGAGATGAGATTGCTGTGTGAAAAAGATCAAGCATGAACATTT 322
Qy 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
```

```
|||||
Db 323 GTGACCTCGAGGACATCTATGAGACACCCACTACTACCTGGTCATGCGATTGTT 382
QY 101 SerGlyGlyLeuPheAspArgLeuGluArgGlyValThrGluLysAspAla 120
Db 383 TCTGGTGGGGAGCTCTTTGACCGGATCCCTGGAGCGGGTGTCTACACAGAACGATGCC 442
QY 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 443 AGTCTGGTATCCAGCAGGTCTTGTGGCAGTGAATACCTACATGAGAATGGCATCGTC 502
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 503 CACAGAGACTTAAAGCCGCAACCTTACCTTACCTTGAAGAGAACTCTTAAGATC 562
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThr-AlaCys 180
Db 563 ATGATCAGTACCTTGGTCTGTCAGATGGAACAGATGGCATGATGCCACCTGCCCTG 622
QY 180 sGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 623 TGGACCCCGAGCTAGCTGGTCCAGAGTGTGGCCAGAAACCTACAGCAAGGCTGT 682
QY 200 lAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPhe-T 220
Db 683 GGATTGCTGCTCCATCGGGTCCATCACCCTACATATTTGCTCTGGGATACCCCGCTCCT 742
QY 220 yrGluGluThrGluSerLysLeuPhe- --GluLysIleLysGluGlyTyrGluPheG 239
Db 743 ATGAAGAAGCGAATCTAAACCTTTTCGAGAAGAATCAAGGAGGCGCTACTATGAGATC 802
QY 239 luSerPro 241
Db 803 TGAGCCCC 810

RESULT 9
BI1818261
LOCUS 603032510F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173587 5',
DEFINITION mRNA sequence.
ACCESSION BI1818261
VERSION BI1818261.1 GI:15928724
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 742)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1432 row: d column: 04
High quality sequence stop: 742.
Location/Qualifiers
1. .742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5173587"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
```

```
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 197 a 180 c 197 g 168 t
ORIGIN
Alignment Scores: 1.16e-105 Length: 742
Pred. No.: 1134.00 Matches: 220
Score: 1134.00 Conservative: 0
Percent Similarity: 99.10% Mismatches: 2
Best Local Similarity: 99.10% Indels: 0
Query Match: 45.13% Gaps: 0
Db: 13
US-09-960-643-2 (1-476) x BI1818261 (1-742)
QY 1 MetGlyArgLysGluGluAspAspCysSerSerTriplyLysGlnThrThrAsnIleArg 20
Db 77 ATGGTCTGAAGGAAGAAGATGATGCTGAGTTCCTGGGAAGAACAGACCACCAACATCCGG 136
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 137 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTTCCTCTGGTG 196
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 197 AAGCAAGACGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCCACCTGCTTC 256
QY 61 ArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
Db 257 CGGCACAGCAGCCTGGAGAAATGAGATTGCTGTGTGAAAAAGATCAAGCATGAAACATT 316
QY 81 ValThrLeuGluAspIleTyrGluSerThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 317 GTGACCTGGAGGACATCTATGAGAGCACCACCACCTACTACCTGGTGCATGCGACTTGT 376
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValThrGluLysAspAla 120
Db 377 TCTGTTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATGCC 436
QY 121 SerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 437 AGTCTGGTATCCAGCAGGCTTCTGCGCAGTGAATACCTACATGAGAATGGCATGCTC 496
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 497 CACAGAGACTTAAAGCCGCAACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 556
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 557 ATGATCACTGACTTTGGTCTGTCCAAGATGGAACAGAATGGCATCATGTCCACTGCTGT 616
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 617 GGGACCCCGAGCTACGGTGTCCAGAAAGTGTGGCCCAAGAAAGCTACAGAGGCTGTG 676
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220
Db 677 GATTGCTGGTCCATCGGGCTCATCACCCTACATATAGTCTCTGGATACCCCGCTTCAT 736
QY 221 GluGlu 222
Db 737 GAAGAA 742
RESULT 10
BI1818261
LOCUS 603032584F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196323 5',
DEFINITION mRNA sequence.
ACCESSION BI1818261
LOCUS 603032584F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196323 5',
DEFINITION mRNA sequence.
ACCESSION BI1818261
```

```
VERSION B1753035.1 GI:15744613
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 731)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: crahps-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUAM1491 row: g column: 12
High quality sequence stop: 729.
FEATURES
Location/Qualifiers
1..731
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5196323"
/lab_host="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
BASE COUNT 190 a 182 c 191 g 168 t
ORIGIN
Alignment Scores:
Pred. No.: 1.93e-104 Length: 731
Score: 1122.00 Matches: 217
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 44.65% Indels: 0
Gaps: 0
DB: 13
US-09-960-643-2 (1-476) x B1753035 (1-731)
QY 2 GlyArgLysGluGluAspAspCysSerSerTriplyLysGlnThrThrAsnIleArgLys 21
Db 73 GGTGGAAGGAGAGAGATGACTGAGTCTCTGGNAGAACACACCAACATCCGGAAA 132
QY 22 ThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuValLys 41
Db 133 ACCTTCAATTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAATTTCTCTGGTGAAG 192
QY 42 GlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPheArg 61
Db 193 CAAGACTGACTGGGAGCTCTTGTCTCTGAAGTCATCAAGAGTCACTGCTGCTCCGG 252
QY 62 AspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIleVal 81
Db 253 GACAGCAGCTGGAGATGAGATGCTGTGTGAAAAGATCAAGCATGAAACATTTGTG 312
QY 82 ThrLeuGluAspIleTyrGluSerThrHisTyrTyrLeuValMetGlnLeuValSer 101
Db 313 ACCCTGGAGGACATCTATGAGAGACCAACCACTACTACCTGCTGATGACAGTGTCT 372
QY 102 GlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAlaSer 121
Db 373 GGTGGGAGCTCTTTGACCGGATCTTGAGCGGGGTCTTACACAGAGAGGATGCCAGT 432
```

```
QY 122 LeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleValHis 141
Db 433 CTGGTGATCCAGCAGGCTCTTCGGCACTGAAATACCTACATGAGAAATGGCATGCTCCAC 492
QY 142 ArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIleMet 161
Db 493 AGAGACTTTAAAGCCCGAAACCTGTTTACCTTACCCCTGAAGAGAACTCTTAAGATCATG 552
QY 162 IleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCysGly 181
Db 553 ATCACTGACTTGGTCTCTCCAGATGGAACAGATGTCATCATGTCCTGCTGCTGGG 612
QY 182 ThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaValAsp 201
Db 613 ACCCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCTACAGCAAGGCTGTGGAT 672
QY 202 CysTTPSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 673 TGCTGTCATCGCGCTCATCACATATATGCTCTGTGGATACCCCCC 723
RESULT 11
BG293660 818 bp mRNA linear EST 21-FEB-2001
LOCUS 602390529F1 NIH_MGC_94 Mus musculus CDNA clone IMAGE:4502479 5',
DEFINITION mRNA sequence.
ACCESSION BG293660
VERSION BG293660.1 GI:13053536
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheraia; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 818)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: crahps-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUAM10371 row: i column: 08
High quality sequence stop: 765.
FEATURES
Location/Qualifiers
1..818
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4502479"
/lab_host="NIH_MGC_94"
/tissue_type="retina"
/note="Organ: eye; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 216 a 191 c 221 g 190 t
ORIGIN
Alignment Scores:
Pred. No.: 2.05e-102 Length: 818
Score: 1103.00 Matches: 228
Percent Similarity: 94.69% Conservative: 4
Best Local Similarity: 93.06% Mismatches: 11
Query Match: 43.89% Indels: 5
Gaps: 0
DB: 12
US-09-960-643-2 (1-476) x BG293660 (1-818)
```


QY 1 MetGlyArgLysGluGluAspSerSerSerTrpLysLysGlnThrThrAsnIleArg 20
|||||
Db 86 ATGGCGGTAAGGAGGAGGAGGAGTCTGCTCTGGAGAAACACAGACCACCAATCAGG 145
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
|||||
Db 146 AAAACCTTCATCTTCATGAAGTCTGGGATCAGGAGCTTTCTCAGAGGTCTTCGGTG 205
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
|||||
Db 206 AAGCAAGAGTGACTGGGAACTCTTTGCTCTGAAATGATATCAAGAGTCACACGCCITC 265
QY 61 ArgAspSerSerLeuGluGlnGluLeuAlaValLeuLysLysIleLysHisGluAsnIle 80
|||||
Db 266 CGGGACAGCAGCCTAGAGATGAGATCGCTGTTGAAAGGATCAAGCATGAGAACATT 325
QY 81 ValThrLeuGluAspIleTrpGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
|||||
Db 326 GTGACCTGGAGGACATCTATGAGACACACCCACTACTACTGCTGTCATGAGCTTGT 385
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
|||||
Db 386 TCTGGAGGTGAGCTTTTACCGGATCTTAGAGCGTGTCTACACAGAAAGGATGCC 445
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
|||||
Db 446 AGCTGTGTCATCCAGCAGGTCTTGTCTGCGGTGAATACCTTCATGAGAAATGGCATCGTC 505
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
|||||
Db 506 CACAGAGACTAAAGCTGAAACCTGCTGTACCTACCCCTGAGGAGAACTCCAGATC 565
QY 161 MetIleThrAspPhe-GlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCy 180
|||||
Db 566 ATGATCAGTCTGTGTGCTATCCAGATGGAGCAGATGGAGTCTATGTC-ACAGCTTG 624
QY 180 sGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVa 200
|||||
Db 625 TGGAGCCCCAGGTACGTGCTCCAGAGTGTCTGGCCCAAGAGCCCTCAGTAAGCTGT 684
QY 200 LysCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro- 220
|||||
Db 685 GGACTGCTGTC-ATTGGTGTATCATACATACATAC- 742
QY 220 YrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluS 240
|||||
Db 743 ATGAAGAAACAGAATCAAGTTTTTCGAGACGATCAGAGGGTTACTACGAAGTTTGAA 802
QY 240 erProPheTrp 243
|||
Db 803 GTTCCGCTCGG 813
RESULT 12
Bg715920 824 bp mRNA linear EST 08-MAY-2001
LOCUS 602676667F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799224 5',
DEFINITION mRNA sequence.
ACCESSION Bg715920
VERSION Bg715920.1 GI:13995107
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 824)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM10688 row: e column: 17
High quality sequence stop: 766.
FEATURES
Location/Qualifiers
1..824
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4799224"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 223 a 194 c 219 g 188 t
ORIGIN

Alignment Scores:
Pred. No.: 3.33e-102 Length: 824
Score: 1101.00 Matches: 222
Percent Similarity: 96.58% Conservative: 4
Best Local Similarity: 94.87% Mismatches: 6
Query Match: 43.81% Indels: 3
DB: 12 Gaps: 0
US-09-960-643-2 (1-476) x Bg715920 (1-824)
QY 1 MetGlyArgLysGluGluAspSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 107 ATGGGTCCAAAGGAAAGAGATGACTGTCAGTCTCTGGAGAAACACAGACCACCATCGG 166
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 167 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAATTTCTCGTGTG 226
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 227 AAGCAAGACTGACTGGGAAGCTCTTCTGCTGAAAGTGCATCAAGAGTCACCTGCCTTC 286
QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
Db 287 CGGGACAGCAGCTGGAGATGAGGTGTCTGTTTGAAGAAAGATCAAGCATGAAACATT 346
QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 347 GTGACCTGGAGGACATCTATGAGACACACCCACTACTACTGCTGTCATGAGCTTGT 406
QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Db 407 TCTGTGGGGAGCTCTTTTACCGGATCATGGAGCGGGGTGTCTACACAGAGAAGATGCC 466
QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 467 AGTCTGGGTATCCACAGCTCTGTGCGCAGTGAATACCTACATGAGAAATGGCATCGTC 526
QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 527 CACAGAGACTTAAACCCCGAAACCTGCTTTTACCTTACCTCCCTGAGAGAACTCTAAGATC 586
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 587 ATGATCACTGACTTTGGTCTGTCCAGATGGAAAGATGGCATCATGCTCCACTGCCTGT 646

Qy 181 GlyThrProGlyTyrVal-AlaProGluValLeuAlaGln-LysProTyrSerLysAlaV 200
 Db 647 GGGACCCAGGCTACGTGGCTGCAGAGTGTGGCCGAGAAACCTACAGCAGGCTG 706
 Qy 200 aAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheT 220
 Db 707 TGGATTCTGGTCCATCGGGGT-ATCACCTAACATATGCTCTGTGGATACCCCCATCT 765
 Qy 220 yrGluGluThrGluSerLysLeuPheGluLysLys 232
 Db 766 ATGAAGAAACGAGCTCTTACGCTTTTCGAAAAATCAAG 803

RESULT 13
 LOCUS BQ949629 918 bp mRNA linear EST 21-AUG-2002
 DEFINITION AGENCOURT_8744314 Lupski_sciatic_nerve Homo sapiens cDNA clone
 IMAGE:6205867 5', mRNA sequence.
 ACCESSION BQ949629
 VERSION BQ949629.1 GI:22365107
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 918)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs.r@mail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM13629 row: g column: 20
 High quality sequence stop: 697.

FEATURES

source
 1..918 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6205867"
 /clone_lib="Lupski_sciatic_nerve"
 /sex="male"
 /tissue_type="sciatic nerve"
 /dev_stage="adult, 70 yr"
 /lab_host="PH10B"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCAGCGCTCG-3' and
 5'-GACTAGTCTAGATCGGAGCGCGCCCTT(15)-3'. Size selected >
 1 kb for average insert length 1.87 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

BASE COUNT 237 a 240 c 250 g 191 t
 ORIGIN

Alignment Scores:

Pred. No.: 3,73e-93 Length: 918
 Score: 1013.50 Matches: 200
 Percent Similarity: 82.86% Conservative: 32
 Best Local Similarity: 71.43% Mismatches: 43
 Query Match: 40.33% Indels: 5
 DB: 14 Gaps: 3

US-09-960-643-2 (1-476) x BQ949629 (1-918)

Qy 1 MetGlyArgLysGluGluAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
 Db 79 ATGGCCCGGAGAAAGCGGAGAGAGAGTCCCTCGTGGAAAAAGAGCTGAAGACATCAAG 138
 Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
 Db 139 AAGATCTTCGAGTTCAAGAGACCCCTCGAACCAGGGGCTTTCCGAAAGTGGTTTGTAGCT 198
 Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAla 59
 Db 199 GAAGAGAAGGCAACTGGCAAGCTCTTGTCTGAAGTGTATCCCTAAGAGGCGCTGAAG 258
 Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
 Db 259 GGCAGGAAAGCAGCATAGAAATGAGATACCCCTCTGAGAAAGATTAAGCATGAAAT 318
 Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
 Db 319 ATTGTGCTCCCTGGAGACATTTATGAAGCCCAATCACCTGTACTTGGTCATCAGCTG 378
 Qy 100 ValSerGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
 Db 379 GTGTCGGTGGAGAGCTTTTACCGGATAGTGGAGAGGGGTTTATACAGAGAGGAT 438
 Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyLe 139
 Db 439 GCCAGCATCTGTATCCGCCAAGTCTTGGACCCGCTGACTATCTCCACAGAAATGGGATC 498
 Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159
 Db 499 GTCCACAGAGACCTCAAGCCCGAAATCTCTTGTACTACAGTCAAGATGAGGAGTCCAAA 558
 Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyLysIleMetSerThr 178
 Db 559 ATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 618
 Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
 Db 619 GCCTGTGGAACCTCCAGGCTATGCTGCTCTGAAGTCTCGCCCAAGAAACCTTACAGCAA 678
 Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuLysGlyTyrProPro 218
 Db 679 GCCGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738
 Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysLysGluGlyTyrTyrGluPhe 238
 Db 739 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 798
 Qy 239 GluSerProPheTrp-AspAspIleSerGluSerAlaLysAspPheIleCysHis-LeuL 258
 Db 799 GACTCTCCCTACTGGGGATGACATCTCCGACTCTGGCAAGAGACTTCATTCGGAACCTGA 858
 Qy 258 euGluLysAspProAsnGluArgTyrThrCysGluLysAlaLysLeuSerHisPro 275
 Db 859 TGGAGAGGAGCCGAAATAAAGATACCCGCTGTGAAGCAGGACGAGCTCGGACACCCC 914

RESULT 14
 LOCUS AW251224 555 bp mRNA linear EST 17-DEC-1999
 DEFINITION UI-R-BJ0-adv-f-07-0-UI.s1 UI-R-BJ0 Rattus norvegicus cDNA clone
 AW251224
 ACCESSION AW251224
 VERSION AW251224.1 GI:6594815
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 555)
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL
MEDLINE
COMMENT

Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Oligo-dT track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA-No.

FEATURES
source

Location/Qualifiers
1..555
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ0-adt-f-07-0-UI"
/clone_lib="UI-R-BJ0"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pMT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; The UI-R-BJ0
library is a subtracted library derived from the UI-R-AA1,
UI-R-AB1, UI-R-AC1, UI-R-AD1, UI-R-AE1, UI-R-AF1, and
UI-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
string of 5-6 nucleotides present between the Not 1 site
and the oligo-dT track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research
6: 791-806, 1996.
TAG_LIB=UI-R-BJ0
TAG_TISSUE=AV canal at 16.5 dpc
TAG_SEQ=GAACC"

BASE COUNT 131 a 134 c 139 g 151 t
ORIGIN

Alignment Scores:

Pred. No.: 1.4e-90 Length: 555
Score: 985.00 Matches: 181
Percent Similarity: 98.91% Conservative: 1
Best Local Similarity: 98.37% Mismatches: 2
Query Match: 39.20% Indels: 0
DB: 10 Gaps: 0

US-09-960-643-2 (1-476) x AW251224 (1-555)

Qy 128 LeuSerAlaValLysTyrLeuHisGluAnGlyIleValHisArgAspLeuLysProGlu 147
|||||
Db 553 TTGTCGCGGTGAAATACCTTCATGAGAAATGGCATGCTTCACAGAGATCAAGGCTGAA 494
Qy 148 AsnLeuLeuTyrLeuThrProGluGluAnSerLysIleMetIleThrAspPheGlyLeu 167
|||||
Db 493 AACCTGCTGACTCACCCCTGAGGAGAACTCCAAGATAATCATCAGTTTGGCCTA 434
Qy 168 SerLysMetGluClnAnGlyIleMetSerThrAlaCysGlyThrProGlyTyrValAla 187
|||||
Db 433 TCCAAGATGGAGCAGAAATGAGTCATGTCACAGCTGTGGGACCCAGGCTATGTGGCT 374
Qy 188 ProGluValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSerIleGlyVal 207
|||||
Db 373 CCAGAAGTGTGGCCCAAGAGCCCTACAGTAAGGCTGTGGACTGCTGCTCCATTGGTGTC 314
Qy 208 IleThrTyrIleLeuLeuCysGlyTyrProPropheTyrGluGluThrGluSerLysLeu 227
|||||
Db 313 ATCACATACATACTGCTGTGGTATCCCTTCTATGAAGAAACAGAAATCAAAAGCTT 254
Qy 228 PheGluLysIleLysGluGlyTyrTyrGluPheGluSerPropheTrpAspAspIleSer 247

Db 253 TTTGAGAAGATCAAGAGGTTACTAGAGTTTGAGTCTCCATCTGGGATGACATTCT 194
Qy 248 GluSerAlaLysAspPheIleCysHisLeuLeuGluLysAspProAnGluArqTyrThr 267
|||||
Db 193 GAGTCAGCCCAAGACTTATTTGCCATCTCTTGAGAGAGGCCCAATGATGACATACACC 134
Qy 268 CysGluLysAlaLeuSerHisProTrpIleAspGlyAsnThrAlaLeuHisArgAspIle 287
|||||
Db 133 TGGGAGAAAGCCCTCAGACACCCCTTGATTTGACGGGAACACAGCCCTGCACCGGACATC 74
Qy 288 TyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTyrTrpArgGlnAla 307
|||||
Db 73 TACCATCTCTCAGCCTCCAGATTTCAGAGAACTTTGCCAAGAGCAAGTGGGAGCAAGCC 14
Qy 308 PheAsnAlaAla 311
|||||
Db 13 TTCATGCGGCC 2
RESULT 15
AL556476
LOCUS AL556476 962 bp mRNA linear EST 16-FEB-2001
DEFINITION AL556476 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DK004YH19 5
prime, mRNA sequence.
ACCESSION AL556476
VERSION AL556476.1 GI:12899184
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 962)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source

Location/Qualifiers
1..962
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DK004YH19"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: Not1; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 218 a 263 c 293 g 185 t
ORIGIN

Alignment Scores:

Pred. No.: 2.66e-88 Length: 962
Score: 966.50 Matches: 189
Percent Similarity: 83.46% Conservative: 38
Best Local Similarity: 69.49% Mismatches: 42
Query Match: 38.46% Indels: 4
DB: 9 Gaps: 3

US-09-960-643-2 (1-476) x AL556476 (1-962)

Qy 12 TrpLysLysGlnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySer 31
|||||
Db 150 TGG---AAGCAGCGGAGGACATTTAGACATCTACGACTTCCGAGATGTTCTGGGCACG 206

```
Qy 32 GlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeu 51
Db 207 GGGGCCCTTCGCGAGGTGATCCTGGCAGAGATAAGAGCAGCAGAGAGCTGGTGGCCATC 266
Qy 52 LysCysIle--LysSerProAlaPheArgAspSerSerLeuGluAsnGluIleAla 70
Db 267 AATGCAATGCCAAGAGGCCCTGGAGGCCAAGGAGGAGCATGGAGAAATGAGATTGCT 326
Qy 71 ValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIleTyrGluSerThr 90
Db 327 GTCTGTCACAAGATCAAGACACCCCAACATGTAGCCCTGGATGACATCTATGAGAGTGGG 386
Qy 91 ThrHisTyrTyrLeuValMetGlnLeuValSerGlyGlyLeuPheAspArgIleLeu 110
Db 387 GGCACCTCTACCTCATCATGTCAGCTGGTGGTGGGGAGGCTCTTTGACCGTATTGTG 446
Qy 111 GluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnGlnValLeuSerAla 130
Db 447 GAAAAAGGCTTCTACACGGAGCGGAGCGCCAGCGGCTCATCTCCAGGTGCTGGATGCT 506
Qy 131 ValLysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLeu 150
Db 507 GTGAATACCTGTCATGACCTGGGCATGTGTACACCGGGATCTCAAGCCAGAGAATCTGCTG 566
Qy 151 TyrLeuThrProGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMet 170
Db 567 TACTACAGCTGGATGAAGACTCCAAATCATGATCTCCGACTTTGGCCCTCTCCAAAGATG 626
Qy 171 GluGlnAsnGly---IleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGlu 189
Db 627 GAGGACCCGGCGAGTGTGCTCCACCGCTGTGGAACTCCGGGATACGTGGCCCTGAA 686
Qy 190 ValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThr 209
Db 687 GTCCTGGCCAGAGCCCTACAGCAAGGCTGTGGATTGGTGTCCATAGGTGTATCGCC 746
Qy 210 TyrIleLeuLeuCysGlyTyrProPheTyrGluGluThrGluSerLysLeuPheGlu 229
Db 747 TACATCTTGCTTGGCGTTACCCYCCCTTCTATGACGAGAAATGATGCCAAACTCTTGAA 806
Qy 230 LysIleLysGluGlyTyrTyrGluPheGluSerProPheTrpAspIleSerGluSer 249
Db 807 CAGATTTTGAAGGCCGAGTACGAGTTTGACTCTCTTACTGGGAGCAGCATCTCTGACTCT 866
Qy 250 AlaLysAspPheIleCysHisLeuLeuGluLysAspProAsnGluArgTyrThrCysGlu 269
Db 867 GCCAAGATTTCATCCGGCACTTKATGGAGAGGA-CCAGAGAAAAGATTACCTGTGAG 925
Qy 270 LysAlaLeuSerHisProTrpIleAspGlyAsnThr 281
Db 926 CAGGCCTTGCAGCACCCCATGGATTGCAGGAGATACA 961
```

Search completed: March 15, 2003, 06:39:01
Job time : 2231 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2003, 05:09:07 ; Search time 113 Seconds
(without alignments)
2956.529 Million cell updates/sec

Title: US-09-960-643-2

Perfect score: 2513
Sequence: 1 MGRKEDDCSWKKQTTNIR.....VKASGSSHRAGQTGVCLIM 476

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-Q-/cgn2.1/USPTO.spool/US09960643/runat.07032003.090518.19918/app_query.fasta_1.647
-DB-PublishedApplications.NA -OFMT=fastap -SUEFTX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX=blowsum62
-TRANS=human40 cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09960643 @cgn_1_1_57.@runat.07032003.090518.19918
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications.NA.*

1:	/cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2.6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3:	/cgn2.6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2.6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2.6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2.6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7:	/cgn2.6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2.6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2.6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10:	/cgn2.6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11:	/cgn2.6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12:	/cgn2.6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13:	/cgn2.6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14:	/cgn2.6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2513	100.0	1738	9 US-09-935-464-4	Sequence 4, Appli
2	2362	94.0	1383	9 US-09-935-464-2	Sequence 2, Appli
3	1250.5	49.8	1772	9 US-10-024-036B-1	Sequence 1, Appli
4	1246	49.6	1074	9 US-10-024-036B-3	Sequence 3, Appli

5	1237	49.2	1578	10 US-09-835-788A-6	Sequence 6, Appli
6	1029	40.9	1372	10 US-09-817-181-1	Sequence 1, Appli
7	865	34.4	157875	9 US-09-935-464-1	Sequence 1, Appli
8	760	30.2	480	9 US-09-935-464-46	Sequence 46, Appli
9	747.5	29.7	1503	10 US-09-797-039-3	Sequence 3, Appli
10	747.5	29.7	2297	10 US-09-797-039-1	Sequence 1, Appli
11	747.5	29.7	3124	12 US-10-153-921-1	Sequence 1, Appli
12	667	26.5	467	9 US-09-935-464-47	Sequence 47, Appli
13	645	25.7	470	9 US-09-935-464-48	Sequence 48, Appli
14	639.5	25.4	3552	9 US-09-934-406-1	Sequence 1, Appli
15	633.5	25.2	474	9 US-09-796-692-7777	Sequence 7777, Ap
16	625.5	24.9	2298	9 US-10-217-357-3	Sequence 3, Appli
17	625.5	24.9	2298	10 US-09-975-326-3	Sequence 3, Appli
18	625.5	24.9	2301	9 US-09-934-406-3	Sequence 3, Appli
19	625.5	24.9	2301	9 US-10-217-357-1	Sequence 1, Appli
20	625.5	24.9	2301	10 US-09-975-326-1	Sequence 1, Appli
21	611.5	24.3	7626	9 US-10-001-835-82	Sequence 82, Appli
22	606.5	24.1	1158	9 US-10-116-332-1	Sequence 1, Appli
23	591.5	23.5	5637	10 US-09-917-800A-1537	Sequence 1537, Ap
24	591	23.5	2824	10 US-09-992-481-5	Sequence 5, Appli
25	587	23.4	3579	9 US-10-024-036B-4	Sequence 4, Appli
26	585.5	23.3	1947	9 US-10-024-036B-6	Sequence 6, Appli
27	585.5	23.3	2454	10 US-09-992-481-3	Sequence 3, Appli
28	584.5	23.3	3705	12 US-10-044-090-227	Sequence 227, App
29	583.5	23.2	1808	10 US-09-740-627-2	Sequence 2, Appli
30	574.5	22.9	2061	12 US-10-096-960-1	Sequence 1, Appli
31	563.5	22.4	1473	9 US-09-938-842A-786	Sequence 786, App
32	562.5	22.4	2735	9 US-10-081-119-15	Sequence 15, Appli
33	556.5	22.1	1833	9 US-09-938-842A-2334	Sequence 2334, Ap
34	549.5	21.9	3061	10 US-09-880-107-2146	Sequence 2146, Ap
35	543	21.6	1488	9 US-09-938-842A-704	Sequence 704, App
36	539.5	21.5	1551	9 US-09-925-299-210	Sequence 210, App
37	539.5	21.5	1551	10 US-09-925-299-210	Sequence 210, App
38	537.5	21.4	1545	10 US-09-922-138-6	Sequence 6, Appli
39	537.5	21.4	1545	10 US-09-841-683-1	Sequence 1, Appli
40	537.5	21.4	2001	10 US-09-841-683-3	Sequence 3, Appli
41	537.5	21.4	2456	10 US-09-922-138-4	Sequence 4, Appli
42	528.5	21.0	1635	9 US-09-938-842A-2344	Sequence 2344, Ap
43	526.5	21.0	1800	9 US-09-938-842A-1063	Sequence 1063, Ap
44	526	20.9	1638	9 US-09-938-842A-1859	Sequence 1859, Ap
45	521	20.7	5926	10 US-09-969-708-302	Sequence 302, App

ALIGNMENTS

RESULT 1
US-09-935-464-4
; Sequence 4, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1738
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-935-464-4

Alignment Scores:
Pred. No.: 6,89e-232 Length: 1738
Score: 2513.00 Matches: 476
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 9 Caps: 0
US-09-960-643-2 (1-476) x US-09-935-464-4 (1-1738)

Qy 1 MetGlyArgLysGluUAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 18 ATGGGTCAAGAGAGAGATGACTGCAGTTCTGGAGAAACAGACCAACATCCGG 77
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 78 AANACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTTCTCAGAAATTTTCTCGTG 137
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 138 AAGCAAGACTGACTGGGAAGCTTTTGCTCTGAAGTGCATCAAGAAGTCACTGCCTTC 197
Qy 61 ArgAspSerSerLeuGluUAsnGluIleAlaValLeuLysLysIleLysHISGluAsnIle 80
Db 198 CGGGACAGCGCTGGAGAAATGAGATTGCTGTGTGAAAGATCAAGCATGAANAACAT 257
Qy 81 ValThrLeuGluUAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
Db 258 GTGACCTTGGAGGACATCTATGAGACACACCCACTACTACCTGGTCATGCAGCTTGT 317
Qy 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
Db 318 TCTGTGGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGTCTACAGAGAAGATGCC 377
Qy 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
Db 378 AGTCGTGTGNTCCAGCAGCTCTTGTGGCAGTGAANTACCTACATGAGATGGCATCGTC 437
Qy 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
Db 438 CACAGAGACTTAAAGCCGAAACCTGCTTTTACCTTACCCCTGAAGAGAACCTTAAGATC 497
Qy 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 498 ATGATCACTGACTTTGGTCTCTCCAAGATGAACAGAAATGCATGTCCACGTCCCTGT 557
Qy 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 558 GGGACCCAGGCTAGTGTGCTCCAGAGTGTGGCCCAAGAACCTTACAGCAAGGCTGTG 617
Qy 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220
Db 618 GATTGCTGGTCCATCGGCGTCACTACCTACATATTGTCTCTGGATACCCCCCATTTCTAT 677
Qy 221 GluGluThrGluSerLysPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Db 678 GAAGAAACGGAGCTTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGATTTGAGTCT 737
Qy 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
Db 738 CCATTTCTGGATGACATTTCTGAGTCAGCCCAAGGACTTTATTTGCCACTTGTGAGAAG 797
Qy 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
Db 798 GATCCGAACGACGGGTACACCTGTGAGAAGCCCTTGAGTCACTCCCTGGATTTGACGGAAAC 857
Qy 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
Db 858 ACAGCCCTTCCACCGGACATCTACCCATCAGTCAGCTCCAGATCCAGAGAACCTTTGCT 917
Qy 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320
Db 918 AAGAGCAAGTGGAGCAAGCCTTCAACGCAGCAGCTGTGGTGCACCACCATGAGGAAGCTA 977
Qy 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProProGluThr 340
Db 978 CACATGAACCTGCACAGCCCGGGCGTCCGCCAGAGGTGGAGAACAGGCCCTGAAACT 1037

Qy 341 GlnAlaSerGluThrSerArgProSerSerProGluIleThrIleThrGluAlaProVal 360
Db 1038 CAAGCCTCAAAACCTCTAGACCCAGCTCCCTGTAGATCACCATCACCAGGACCTGTC 1097
Qy 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
Db 1098 CTGAGCACAGTGTAGCACTCCTGCCCTGACCCATTACCTGCCAGCATGGCCGCCGG 1157
Qy 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
Db 1158 CCCACTGCCCTGGTGGCAGTCCCTCAACTGCTGCTCAATGGCTCCCTCCACATCAGC 1217
Qy 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysCysSer 420
Db 1218 AGCAGCTGGTGGCCATGATCATGAGGGTCCCTGGCCGCCGGCCCTGGCTGCTCTCC 1277
Qy 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu 440
Db 1278 AGCTGCTGACATTTGGAGCAAGGAAGTCTCTCTACTGCTCTGAGCCACACTCCTC 1337
Qy 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460
Db 1338 AAAAAGGCCAAACAAAAACAGAACTTCAAGTCGGAGGTCATGGTACCAGTTAAAGCAGT 1397
Qy 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
Db 1398 GGCAGCTCCCACTGCCGGCAGGGCAGACTGGAGTCTGTCTATTATG 1445

RESULT 2
US-09-935-464-2
; Sequence 2, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH
; FILE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/14702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-2

Alignment Scores:
Pred. No.: 1,55e-217 Length: 1383
Score: 2362.00 Matches: 447
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 93.99% Indels: 0
DB: 9 Gaps: 0

US-09-960-643-2 (1-476) x US-09-935-464-2 (1-1383)

Qy 1 MetGlyArgLysGluUAspAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 1 ATGGGTCAAGAGAGAGATGACTGCAGTTCTGGAGAAACAGACCAACATCCGG 60
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 61 AANACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTTCTCAGAAATTTTCTCGTG 120
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 121 AAGCAAGACTGACTGGGAAGCTTTTGCTCTGAAGTGCATCAAGAAGTCACTGCCTTC 180


```
Db 289 GAAGAGAAGCAACTGGCAAGCTCTTTGCTGTGAAGTGTATCCCTAAGAAGCGCGCTGAAG 348
Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysLeuHisGluAsn 79
Db 349 GGCACGAAGAAGCAGATAGAGATAGATAGCGCTCGTGGAGAAGATTAAAGCATGAATAAT 408
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
Db 409 ATTGTTGCCCTGGAGACATTTATGAAGCCCAATCACTTACTGTGTCATGCAGCTG 468
Qy 100 valSerGlyGlyLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysasp 119
Db 469 GTCTCGGTGGAGAGCTGTTGACCGATAGTGGAGAGGGTTTTATACAGAAGGAT 528
Qy 120 AlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
Db 529 GCAGACACTGTATCCGCCAAGTCTTGGAGCGGTGTACTATCTCCACAGAATGGGCATC 588
Qy 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluLysSerLys 159
Db 589 GTCCACAGAGACCTCAAGCCGCAAAATCTTGTACTACAGTCAAGATGAGGAGTCCAAA 648
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly---IleMetSerThr 178
Db 649 ATAATGATCAGTGACTTTGGATTGTCAAAAATGGAGGCCAAAGGAGATGTGATGTCCACT 708
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
Db 709 GCCTGTGGAACTCCAGGCTATGTCTGCTGAAAGTCTGCCCGCCAGAAACCTTACAGCAAA 768
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 769 GCGGTTGACTGCTGGTCCATCGAGTGAU TGCTTACATCTTGCTCTGCGCTACCTCTCT 828
Qy 219 PheTyrGluCluThrGluSerLysLeuPheGluLysIleLysGluCluTyrTyrGluPhe 238
Db 829 TTTTATGATGAANAATCAAGCTCTTTGAGCAGTCAAGCGGCAAGATATGAGTTT 888
Qy 239 GluSerProPheTrpAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 889 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAGAAAGACTTCAATTCGGAACCTGATG 948
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleasp 278
Db 949 GAGAAGACCCGAAATAAAGATACACGTGTGACAGCAGCTCGGACCCATCGATCGCT 1008
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
Db 1009 GGTGACACAGCCCTCAACAAAACATCCAGAGTCCGTGAGCGCCAGATCCGGAACAC 1068
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
Db 1069 TTTGCCAAGACAAATGGAGACAAGCATTTAATGCCACGCGCTCGTGAGACATATGAGA 1128
Qy 319 LysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArg----- 336
Db 1129 AAACCTACACTCGCGCAGCAGCCCTGGACAGTTCAAAATGCAAGTGTTCGAGCAGCCCTCAGT 1188
Qy 337 ---ProProGluThrGlnAlaSerGluThrSerArgProSerSerProGluIleThrIle 355
Db 1189 TTGGCCAGCCAAAAGACTGTGGTATGTAGCAAAACAGATC-CCTCAGCTGACACTG 1247
Qy 356 ThrGluAlaProValLeuAspHisSerValAlaLeuProAla----- 369
Db 1248 AA-GACGAGCTGGGTGGAGAGGAGCGCGCATCTGCCGAGCAGCCTCCTGTGTTGCTC 1306
Qy 370 -----LeuThrGlnLeuProCys----- 375
Db 1307 AGCGGCTTTCTATACTAATATCCATGTGATGCGACCCCTAGGACTTTTTTAACATGTAAT 1366
Qy 376 -----GlnHis---GlyArg----- 379
Db 1367 CACTGGGCGGGTGCAGTGCTCACGCTCTAATCCCAACACTTTCGGAGGCTGAGGCAG 1426
```

```
Qy 380 -----ArgProThrAlaPro----- 384
Db 1427 GAGGACTGTTTGTAGTTCAGGAGTTTTTAAGACCAAGCTGACCAACATGTTGAAACCCCATC 1486
Qy 395 -----GlyGlyArgSerLeuAsnCysLeuValAsn 394
Db 1487 TCTACTAAATATAAAAATTAGCGGGTGTGGTGGCAGCACCTGTATATGTACGTACTT 1546
Qy 395 GlySerLeuHisIleSerSerSerLeuValProMetHisGlnGlySerLeuAlaAlaGly 414
Db 1547 GGGAGGCTGAGGCGAGGAGAACTTGAACCA-----GGA 1582
Qy 415 ProCysGlyCysCysSerSerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCys 434
Db 1583 AGCGGAGGTTCG-----AATGAGCTGAGATCACCACCTGC 1618
Qy 435 SerGlu-----ProThrLeuLeuLysLysAlaAsnLysLysGlnAsn 448
Db 1619 ACTCCAGCTGGGTGACAGATTGAGACTCCCTCTCAAAAAAAGGAAATCATTTGAAC 1678
Qy 449 PheLysSerGluValMetValProValLysAlaSerGly 461
Db 1679 ACTCGTGAACCTAGGTATTGCTATATTCATTACGGT 1717
RESULT 4
US-10-024-036B-3
; Sequence 3, Application US/10024036B
; Publication No. US20030028004A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 68730 and 69112, Protein Kinase
; FILE OF INVENTION: Molecules and Uses Therefor
; FILE REFERENCE: MPI2000-521PIR(M)
; CURRENT APPLICATION NUMBER: US/10/024,036B
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/258222
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-024-036B-3
Alignment Scores:
Pred. No.: 1.59e-110 Length: 1074
Score: 1246.00 Matches: 237
Percent Similarity: 85.49% Conservative: 40
Best Local Similarity: 73.15% Mismatches: 45
Query Match: 49.58% Indels: 2
DB: Gaps: 2
US-09-960-643-2 (1-476) x US-10-024-036B-3 (1-1074)
Qy 1 MetGlyArgLysGluLysAspCysSerSerTrpLysLysGlnThrThrAsnIleArg 20
Db 1 ATGGCCCGGAGAGCGCGAGCAGCTCCTCTGGAAGAGCAAGCTGGAAGACATCAAG 60
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 61 AGATCTTCGAGTCTCAAGAGACCCCTCGAAGCCGGGGCTTTTCGGAAGTGGTTTACGT 120
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle---LysLysSerProAla 59
Db 121 GAAGAGAAGCACTGGCAAGCTCTTGTGTGAAGTGTATCCCTAAGAAGCGCGTGAAG 180
Qy 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 181 GCAAGAGAAGCAGATAGAGATAGCGCTCTGAGAAAGATTAAAGCATGAAAT 240
Qy 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
```



```
Db 241 ATTGTTCCCTGGAGACATTTATGAAGCCCAAAATCACCCTGCTACTGGTTCATGCAGCTG 300
Qy 100 ValSerGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
Db 301 GTGTCGGTGGAGAGCTGTTGACCGGATAGTGGAGAGGGGTTTATACAGAGAAGAT 360
Qy 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
Db 361 GCCAGCACTCTGATCCGCCAAGCTTTGGACGCGCTGTACTATCTCCACAGAAATGGGCATC 420
Qy 140 ValHisArgAspLeuLysProGluAsnLeuTyrLeuThrProGluGluAsnSerLys 159
Db 421 GTCCACAGACCTCAAGCCCGAAATCTCTGTACTACAGTCAAGATGAGGAGTCCAAA 480
Qy 160 IleMetIleThrAspPheGlyLeuSerLysMetGluLeuAsnGly---IleMetSerThr 178
Db 481 ATAATGATCAGTGACTTTGGATTGTCAAAAATGGAGGGCAAGAGATGTGATGTCCACT 540
Qy 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
Db 541 GCCTGTGGAACCTCAGGCTATGCGTCCTGAAAGTCTCGCCAGAAACCTTACAGCAA 600
Qy 199 AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro 218
Db 601 GCCCTTGACTGCTGCTCCAGGAGTGTGCTCATCTTGTCTCGCGCTACCTCCT 660
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
Db 661 TTTTATGATGAAATGACTCAAGCTCTTTTGAGCAGATCTCTCAAGGCGGAATATGAGTTT 720
Qy 239 GluSerProPheTrpAspPheSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 721 GACTCTCCCTACTGGGATGACATCTCGACTCTGCAAAAGACTTCATTGCGAACCTGATG 780
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
Db 781 GAGAAGGACCGCAATAAAGATACACGTGTGAGCAGCAGCTCGCGCACCCATGGATCGCT 840
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
Db 841 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCCGTCAGCGCCAGATCCGGAAAAAC 900
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisMetArg 318
Db 901 TTTGCCAAGAGCAATGAGACAAAGCATTTAATGCCACGCCCTGCTGTGACACATATGAGA 960
Qy 319 LysLeuHisMet 322
Db 961 AAACCTACACCTC 972
```

RESULT 5

```
US-09-835-788A-6
; Sequence 6, Application US/09835788A
; Patent No. US20020077458A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, a
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: P018P1
; CURRENT APPLICATION NUMBER: US/09/835,788A
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US00/28666
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/159,585
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/167,246
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1578
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-835-788A-6
```

Alignment Scores:

```
Pred. No.: 2,07e-109 Length: 1578
Score: 1237.00 Matches: 235
Percent Similarity: 86.16% Conservativeness: 39
Best Local Similarity: 73.90% Mismatches: 42
Query Match: 49.22% Indels: 2
DB: 10 Gaps: 2
```

US-09-960-643-2 (1-476) x US-09-835-788A-6 (1-1578)

```
Qy 7 AspAspCysSerSerTrpLysGlnThrThrAsnIleArgLysThrPheIlePheMet 26
Db 15 GAGAGCAGCTCTCTCTCGAAAAAGCAAGCTGAAGACATCAAGAAAGATCTTCGAGTTCAAA 74
Qy 27 GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly 46
Db 75 GAGACCTCGGNAACCGGGGCGCTTTCCAGAGTGGTTTAGCTGAAGAGAGGCAAGCTGGC 134
Qy 47 LysLeuPheAlaLeuLysCysIle---LysLysSerProAlaPheArgAspSerSerLeu 65
Db 135 AAGCTCTTTGCTGTGAAGTGTATCCTTAAGAGCGCTGAAGGCAAGGAAAGCAGCAT 194
Qy 56 GluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAsp 85
Db 195 GAGAATGAGATAGCGCTCTGAGAAAGATTAAAGCATGAAATATTTGTTCCTCGAAGAC 254
Qy 86 IleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuValSerGlyGlyGluLeu 105
Db 255 ATTTATGAAGGCCCAAAATCACCTGTACTTGGTCATGACAGCTGGTCTCGGTGGAGAGCTG 314
Qy 106 PheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGln 125
Db 315 TTTGACCGGATAGTGGAGAGGGGTTTATACAGAGAGGATGCGACACTCTGTATCGCG 374
Qy 126 GlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLys 145
Db 375 CAAGTCTTGGAGCGGTGTACTATCTCCACAGATGGGCATCGTCCACAGAGACTCAAG 434
Qy 146 ProGluAsnLeuLeuTyrLeuThrProGluAsnSerLysIleMetIleThrAspPhe 165
Db 435 CCCGAAATCTCTTCTACTACAGTCAAGATGAGGAGTCCAAAATAATGATCAGTGACTTT 494
Qy 166 GlyLeuSerLysMetGluGlnAsnGly---IleMetSerThrAlaCysGlyThrProGly 184
Db 495 GGATTTGTCAAAAATGGAGGCAAGAGATGTGATGTCCACTGCTGTGGAACTCCAGGC 554
Qy 185 TyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSer 204
Db 555 TATGTCGCTCTGAAAGTCTCTGCCCCAGAAACCTTACACAAAGCGGTGACTGCTGTGTC 614
Qy 205 IleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyrGluGluThrGlu 224
Db 615 ATCGGAGTGATTGCTCTACATCTTGCTGCGGCTACCTCTCTTTTATGATGATAAATGAC 674
Qy 225 SerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSerProPheTrpAsp 244
Db 675 TCCAAGCTCTTTTGAGCAGATCTCTCAAGGCGGAATATGAGTTTGACTCTCTCCCTACTGG 734
Qy 245 AspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLysAspProAsnGlu 264
Db 735 GACATCTCCGACTCTGCAAAAGACTTCTCGGAACCTGATGGAGAGAGGCCCAAAATAA 794
Qy 265 ArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsnThrAlaLeuHis 284
Db 795 AGATACAGTGTGAGCAGGAGCTCGGCACCCATGGATCGCTGGTGACACAGCCCTCAAC 854
Qy 285 ArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrp 304
Db 855 AAAACATCCACGAGTCCGTCAGGCCCCAGATCGGAAAAAACTTTGGCAAGAGCAAAATGG 914
```

```
QY 305 ArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeuHisMet 322
|||||
Db 915 AGACAAGCAATTAATGCGCCAGCGCGCTGCGTGCACATATGAGAAAACCTACACCTC 968

RESULT 6
US-09-817-181-1
; Sequence 1, Application US/09817181
; Patent No. US20020142427A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001189
; CURRENT APPLICATION NUMBER: US/09/817,181
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Human
US-09-817-181-1

Alignment Scores:
Pred. No.: 1.59e-89 Length: 1372
Score: 1029.00 Matches: 224
Percent Similarity: 64.20% Conservative: 45
Best Local Similarity: 53.46% Mismatches: 90
Query Match: 40.95% Indels: 60
DB: 10 Gaps: 9

US-09-960-643-2 (1-476) x US-09-817-181-1 (1-1372)

QY 60 PheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
|||||
Db 198 TTCGAGAT-----GTTCTGGGCGACGATCAAGCACCCCAAC 233

QY 80 IleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeu 99
|||||
Db 234 ATTGAGCCCTGGATGACATCTATGAGAGTGGGGGCCACCTCTACCTCATCATGAGCTG 293

QY 100 ValSerGlyGlyLeuPhePheArgIleLeuGluArgGlyValTyrThrGluLysAsp 119
|||||
Db 294 GTGTGGGTGGGGAGCTCTTTCACCGTATTGTGAAAAAGCGCTTACACGGAGCGGAC 353

QY 120 AlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIle 139
|||||
Db 354 GCCAGCGGCTCATCTTCCAGGTGCTGGATGCTGTGAATACCTGCATGACCTGGCAT 413

QY 140 ValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLys 159
|||||
Db 414 GTACACCGGGATCTCAAGCAGAGAATCTGCTGTACTACAGCGCTGGATGAAGACTCCAAA 473

QY 160 IleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly--IleMetSerThr 178
|||||
Db 474 ATCATGATCTCCGACTTTGGCCTCTCCAGATGGAGGACCCGGGCGAGTGTCTCCACC 533

QY 179 AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys 198
|||||
Db 534 GCCTGTGGAACTCCGGGATACGTGGCCCTGAAGTCTGCTCCGCCAAGCCCTACAGCAAG 593

QY 199 AlaValAspCysTrpSerIleGlyValIleThrThrIleLeuLeuCysGlyTyrProPro 218
|||||
Db 594 GCTGTGGATTGTGTCTCCATAGTGTTCATCGCCTACATCTTGTCTGTGGGTTTACCTCCC 653

QY 219 PheTyrGluGluThrCysLeuSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
|||||
Db 654 TTCTATGACGAGAATGATGCCAACTCTTTGAACAGATTTGAAGGCGCGAGTACAGATT 713

QY 239 GluSerProPheTrpAspPheSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
|||||
Db 714 GACTCTCCTTACTGGGACGACATCTCTGACTCTGCCAAGATTTCATCCGGCATTGATG 773
```

```
QY 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrPilleAsp 278
|||||
Db 774 GAGAAGACCAGAGAAAGATTACCTGTGACGAGSCCTTGACAGCACCCTGATGATGCA 833

QY 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
|||||
Db 834 GGAGATACAGCTCTAGATAAGAATATCCACCATGTCGCTGAGTGCAGCATCAAGAAGAAC 893

QY 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArg 318
|||||
Db 894 TTTCGCAAGACAGTGAAGACAGCTTCAATGCCACGGCTGTGTGTCGCGCACATGAGG 953

QY 319 LysLeuHisMetAsnLeuHisSerProGly-----Val 329
|||||
Db 954 AAACCTGCAGCTGGGACACGAGGAGGCGAGGCGAGCGAGCCATGGGAGCTG 1013

QY 330 ArgProGluValGluAsnArgProProGluThrGlnAlaSerGluThrSerArgProSer 349
|||||
Db 1014 CTGACACCATGCTGGCTGGGGGGCG-----GCAGCTGGCTGTGCTCTCGAGACTGC 1064

QY 350 SerProGluIleThrIleThrGluAlaProValLeuAspHisSerValAlaLeuProAla 369
|||||
Db 1065 TGCCTGGAGCGGGGCACAGAACTGTCCCCACACTGCCCCACACGCTC----- 1112

QY 370 LeuThrGlnLeuProCysGlnHisGlyArgArgProThrAlaProGlyGlyArgSerLeu 389
|||||
Db 1113 -----TAGGGCCCTGGACCTCGGTCATG 1136

QY 390 AsnCysLeuValAsnGlySerLeuHisIleSerSerSerLeuValProMetHis----- 407
|||||
Db 1137 ATCTCTCGCTGGGAGG-----CTTGGGGGACGCTGTCCCTTCCCTCCCTCGCTGA 1187

QY 408 GlnGlySerLeuAlaAla----- 413
|||||
Db 1188 ACCGGAGTTCTCGCCTCTCCCTCCCTCACCTGCTTCCCTACCACCTCCTCACTGCAT 1247

QY 414 GlyProCysGlyCysCysSerSerCysLeuAsnIleGlySerLysGlyLysSerSerTyr 433
|||||
Db 1248 TTTCCATACAAATGTTTCTATTTTATTGTTCTTCTTGTATTAAGGAAG----- 1298

QY 434 CysSerGluProThrLeuLysAlaAsnLysGlnAsnPheLysSerGlu 452
|||||
Db 1299 ---ATAAACCATCTTAAAAA----- 1352

RESULT 7
US-09-935-464-1
; Sequence 1, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 157875
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-1

Alignment Scores:
Pred. No.: 1.15e-70 Length: 157875
Score: 865.00 Matches: 202
Percent Similarity: 42.44% Conservative: 0
Best Local Similarity: 42.44% Mismatches: 3
```

Query Match: 34.42% Indels: 272
DB: 9 Gaps: 2
US-09-960-643-2 (1-476) x US-09-935-464-1 (1-157875)
Qy 272 LeuSerHisProTrp-IleAspGlyAsnThrAlaLeuHisArgAspIleTyProSerVa 291
Db 145909 CTCCTACATCCACAGGATTACAGGAAACACAGCCCTCCACCGGACATCTACCCATCAGT 145968
Qy 291 lSerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrpArg----- 305
Db 145969 CAGCTCCAGATCCAGAGAACTTTCCTAAGACCAAGTGGAG-GGTAAGCTGTCTCTCTCC 146027
Qy 305 ----- 305
Db 146028 AGGGGTGGGAAAGCTGTTCTGGGCCCTTGGAGGCTGGCTGGCAGGGCTGACATNAGG 146087
Qy 305 ----- 305
Db 146088 GCTTTCCTTGGGATGTCCAGAAAGGCAATTGGAGCTCCGTGTACCCCTCTCTGAAATGAGAA 146147
Qy 305 ----- 305
Db 146148 GTGGCACCCAGGTTTCAAGAGGCCACAGGCCAAAGGGAAAGTTTAAGCTCCCAAGGCCCT 146207
Qy 306 -----GlnAlaPheAsnAlaAl 311
Db 146208 CTGAGTTGCAAGAGCCAGAGGCTGCTCTGTGTCTCTCTAGCAAGCCCTTCAAGCGAGC 146267
Qy 311 aAlaValValHisHisMetArgLysLeuHisMetAsnLeuHisSerProGlyValArgPr 331
Db 146268 ACCTGTGTGCACCATAGGAGGCTACACATGAACCTGCACAGCCCGGGCTCCGCC 146327
Qy 331 oGluValIleLysAsnArgProGluThrGlnAlaSerGluThrSerArgProSerSerPr 351
Db 146328 AGAGTGGAGAACAGCGCGCTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTCCCC 146387
Qy 351 oGluIleThrIleThrGluAlaProValLeuAspHisSerValAlaLeuProAlaLeuTh 371
Db 146388 TGAGATCACCATCACCAGGACACTGTCTGGACACAGTGTAGCACTCCCTGCCCTTGAC 146447
Qy 371 rGlnLeuProCysGlnHisGlyArgArgProThrAlaProGlyArgSerLeuAsnCy 391
Db 146448 CCAATTACCTGCCAGCATGGCCGCGGCCACTGCCCTGGTGGCAGGTCCCTCAACTG 146507
Qy 391 sLeuValAsnGlySerLeuHisIleSerSerSerLeuValProMetHisGlnGlySerLe 411
Db 146508 CCTGGTCAATGGCTCCCTCCACATCAGCAGCAGCCTGGTGGCCATGCATCAGGGGTCCCT 146567
Qy 411 uAlaAlaGlyProCysGlyCysCysSerSerCysLeuAsnIleGlySerLysGlyLysSe 431
Db 146568 GGCCGCGGGCCCTGTGGCTGTCTCCAGCTGCCTCCAGCTGCCTGGAATTTGGGAGCAAGGAAGTC 146627
Qy 431 rSerTyrcysSerGluProThrLeuLeuLysLysAlaAsnLysLysGln----- 447
Db 146628 CTCCTACTGCTCTGAGCCACACTCTCCAAAAAGGCCAACAAAAACAGTACGTATTTT 146687
Qy 447 ----- 447
Db 146688 AGCCAAAGATGGAGCCCGCAGCTTGGGTCTGAAAGAAATCGGTCAACAGGACTGAAAGAAA 146747
Qy 447 ----- 447
Db 146748 TGGACACAAAGCCCTCTCCACTCATAGCAGCTATATAGGGAGGATGAGTCTCGATG 146807
Qy 447 ----- 447
Db 146808 AGGGGCAAGGAAATGCTTCCAGCCCTGTCCCATCATCTACTAGTTGTGAGGCCATTTG 146867
Qy 447 ----- 447
Db 146868 ACGAGTCAATGCTCTGCTTCTATTTTGCTCATCTGTAAACTGGGTGAAGTCTATTATAC 146927

Qy 447 ----- 447
Db 146928 CCAATTAGCTGCTCACAGAGGTACTATAAAAGTTAGTGAAGCAACAGAGATGATAAA 146987
Qy 447 ----- 447
Db 146988 AGCCCAAGAAACTGGGTATTGTCGAAGTGGTGCAGCATCAGATTCCCCGAGTTTCCAGGG 147047
Qy 447 ----- 447
Db 147048 CCTCTGTGCTTTGCAGAAGCAAGTGTAGGGGGCTTGGTTATCTTTATCTTTTCTCAGG 147107
Qy 447 ----- 447
Db 147108 TTCTTGAACCTTCTGAACCTCATCAGGGGAGAGCTGAGTGGATGCCACAGGCACAGAAA 147167
Qy 447 ----- 447
Db 147168 CTACCACCTCTGCCCTGCCCTGCCCTGACTCTTCTTCCCTCAATACTTCAAGGTTGTTGC 147227
Qy 448 -----AsnPhelYsSerGluValMetValProValLysAlaSerGlyS 462
Db 147228 TTCAATTTCTTCTAGGAAGTTCAAAGTCGGAGGTCTACCGTTAAAGCCAGTGGCA 147287
Qy 462 erSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
Db 147288 GCTCCCACTGCCGGGCGAGGCGAGACTGGAGTCTGTCTCAATTATG 147331
RESULT 8
US-09-935-464-46
; Sequence 46, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: n-a or g or c or t/u
US-09-935-464-46

Alignment Scores:
Pred. No.: 2,22e-64 Length: 480
Score: 760.00 Matches: 150
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.24% Indels: 0
DB: 9 Gaps: 0

US-09-960-643-2 (1-476) x US-09-935-464-46 (1-480)

Qy 1 MetGlyArgLysGluAspCysSerSerTrpLysLysGlnThrAsnIleArg 20
Db 30 ATGGGTCGAAAGGAGAAAGATGACTGCAGTTCCTGGAGAAACAGACACCAACATCCGG 89
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 90 AAAACCTTCATTTTATGAAAGTGTGGATCAGGAGCTTCTCAGAAGTTTCTCTGGTG 149
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60

	Db	150	AAGCAAGACTGACTGGGAAGCTCTTTGCTGTGAAGTGCATCAAGAAGTCACCTGCCTTC	209
	Qy	61	ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle	80
	Db	210	CGGACAGCACCTGGAGAATGAGATTGCTGTGTTGAAANAAGATCAAGCATGNAACATT	269
	Qy	81	ValThrLeuGluAspIleTy rGluSerThrThrHisTy rTy rLeuValMetGlnLeuVal	100
	Db	270	GTGACCCCTGGAGGACATCTATGAGACGCCACCACCTACTACCTGTGTCATGCAGCTTGTT	329
	Qy	101	SerGlyClyGlutLeupheAspArgIleLeuGluUargGlyValTy rThrGluLysaspAla	120
	Db	330	TCTGTGTGGGAGCTCTTTGACCGGATCCCTTGAGCGGGGTCTTACACAGAGAAGGATGCC	389
	Qy	121	SerLeuValIleGlnGlnValLeuSerAlaValLysTy rLeuHisGluAsnGlyIleVal	140
	Db	390	AGTCTGTGTATCCAGCANGTCTTGTGCGCNGTGAATACCTACNTGAGAATGGCATCGTC	449
	Qy	141	HisArgAspLeuLysProGluAsnLeuLeu	150
	Db	450	CACAGAGACTTAAGGCCGAAAACCTTGCTT	479

RESULT 9

```

US-09-797-039-3
; Sequence 3, Application US/09797039
; Patent No. US20020042099A1
; GENERAL INFORMATION:
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 2504, 15377, and 14760, NOVEL PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 10448-017001
; CURRENT APPLICATION NUMBER: US/09/797,039
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-797-039-3

```

Alignment Scores:	
Pred. No.:	1,97e-62
Score:	747.50
Percent Similarity:	53.29%
Best Local Similarity:	35.39%
Query Match:	29.75%
DA:	10
Length:	1503
Matches:	172
Conservative:	87
Mismatches:	175
Indels:	53
Gaps:	9

US-09-960-643-2 (1-476) x US-09-797-039-3 (1-1503)

Qy	15	GlnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySerGlyAlaPhe	34
Db	46	CAGCCATCGGAGGTGACTGCATGATGTAU	105
Qy	35	SerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle	54
Db	106	TGTCAAATCTTCGGGGCAAGGACAAGACACAGGCAAGCTGCACACCTCCAGAAAGTTC	165
Qy	55	LysLysSerProAlaPheArgAsp-----SerSerLeuGluAsnGluIle	69
Db	166	CRGAAG-----CGGACGCGCCGCAAGGTGCGGAAAGCTGCCAAGACAGAGATA	213
Qy	70	AlaValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIleTyrGluSer	89
Db	214	GGCATCTCAAGATGGTGAAGCATCCCAACATCCTACAGCTGGTGGATGTGTTTGTGACC	273
Qy	90	ThrThrHisTyrTyrLeuValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIle	109

```

Qy 449 eLysSerGluValMetVal-----ProValLysAlaSerGlySerHisCysAr 466
      :::::      |||      |||      |||
Db 1356 GGCACCAAGCAGCAGCTGCCACCCCTGAGCGCGCTATGCGCCAGCGGACAGCAGCCCC 1415
      |||      |||      |||      |||
Qy 466 gAlaGlyGlnThrGly 471
      |||      |||||
Db 1416 AGAGGGCGCCACAGGC 1431
      |||      |||||

RESULT 10
US-09-797-039-1
: Sequence 1, Application US/09797039
: Patent No. US20020042099A1
: GENERAL INFORMATION:
: APPLICANT: Olandt, Peter J.
: APPLICATOR: Kapeller-Libermann, Rosana
: TITLE OF INVENTION: 2504, 15977, AND 14760, NOVEL PROTEIN
: DATE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
: FILE REFERENCE: 10448-017001
: CURRENT APPLICATION NUMBER: US/09/797,039
: CURRENT FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: US 60/186,061
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 2297
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (154)...(1656)
: NAME/KEY: misc_feature
: LOCATION: (1)...(2297)
: OTHER INFORMATION: n = A,T,C or G
US-09-797-039-1

Alignment Scores:
Pred. No.: 3,74e-62 Length: 2297
Score: 747.50 Matches: 172
Percent Similarity: 53.29% Conservative: 87
Best Local Similarity: 35.39% Mismatches: 175
Query Match: 29.75% Indels: 53
DB: 10 Gaps: 9

US-09-560-643-2 (1-476) x US-09-797-039-1 (1-2297)

Qy 15 GlnThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySerGlyAlaPhe 34
      |||      ::      ::      ::      ::      ::      ::
Db 199 CAGCCATCGAGGTGACTCACAGATATGTTGGACAGGTCATCAAGACTGAGGAGTTT 258
      |||      ::      ::      ::      ::      ::      ::
Qy 35 SerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle 54
      |||      |||      |||      |||      |||      |||      |||
Db 259 TGTGAATCTTCGGGCCAAGCAAGCAGCAGCAGCAAGCTGCACACCTGCCAAGAGTTC 318
      |||      |||      |||      |||      |||      |||      |||
Qy 55 LysLysSerProAlaPheArgAsp-----SerSerLeuGluAsnGluIle 69
      ::      |||      |||      |||      |||      |||      |||
Db 319 CAGAAG-----CGGCACGCCCGCAAGTGCAGAACTGCCACAGACAGAGNTA 366
      ::      |||      |||      |||      |||      |||      |||
Qy 70 AlaValLeuLysIleLysHisGluAsnIleValThrLeuGluAspIleTyrGluSer 89
      ::      |||      |||      |||      |||      |||      |||
Db 367 GGCATCTCAAGATGTGTGAAGCATCCCAACATCTTACAGCTGTTGGATGTGTTGTGACC 426
      ::      |||      |||      |||      |||      |||      |||
Qy 90 ThrThrHisTyrTyrLeuValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIle 109
      |||      |||      |||      |||      |||      |||      |||
Db 427 CGCAAGAGTACTTTATCTCTCTGAGCTGGCCAGCGGGAGGAGGTGTTTGACTGGATC 486
      |||      |||      |||      |||      |||      |||      |||
Qy 110 LeuGluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnValLeuSer 129
      |||      |||      |||      |||      |||      |||      |||
Db 487 CTGGACCAAGGCTACTACTCTCGAGCGAGCAGACAGCAACGCTGGTACGGCAAGTCTTGAG 546
      |||      |||      |||      |||      |||      |||      |||
Qy 130 AlaValTyrTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeu 149
      |||      |||      |||      |||      |||      |||      |||

```

```

RESULT 11
US-10-153-921-1
; Sequence 1, Application US/10153921
; Patent No. US20020142430A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO00612DIV
; CURRENT APPLICATION NUMBER: US/10/153,921
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/207,281
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 09/734,030
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3124
; TYPE: DNA
; ORGANISM: HOMO SAPIEN
US-10-153-921-1

Alignment Scores:
Pred. No.:      5,95e-62          Length:      3124
Score:           747.50           Matches:     177
Percent Similarity: 54.91%       Conservative: 80
Best Local Similarity: 37.82%    Mismatches:  150
Query Match:      29.75%        Indels:      65
DB:                12            Gaps:         10

US-09-960-643-2 (1-476) x US-10-153-921-1 (1-3124)
Qy   15 GLnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySerGlyAlapHe 34
Db   256 CAGCCATCGGAGGTGGCTACAGCATGATTGTTGGCAGAGTGCATCAAGACATGAGAGATT 315
Qy   35 SerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle 54
Db   316 TGTGAATCTTCGCGGCCAACAGCACAGCAGCGCAAGTGCCACTGTCACAAGAAGTTC 375
Qy   55 LysLysSerProAlaPheArgASP-----SerSerLeuGluAsnGluIle 69
Db   376 CAGAAG-----CGGAGCGCGCCAAAGTGCAGAAAGTGCACAAAGACGAGATA 423
Qy   70 AlaValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIleTyTrGluSer 89
Db   424 GCATCTCAAGATGGTGCAGCATCCCCAACATCTACAGCTGGTGGATGTGTGTGTGACC 483
Qy   90 ThrThrHisTyTrLeuValMetGlnLeuValSerGlyGlyGlyLeuPheAspArgIle 109
Db   484 CGCAAGGAGTAGTATTATCTTCCTGGAGTGCACACGGGAGGAGGTGTTTACTGGATC 543
Qy   110 LeuGluArgGlyValTyThrGluLysaspAlaSerLeuValIleGlnGluValLeuSer 129
Db   544 CTGACACGAGGCTACTACTCGGACGACACACAAAGCAACGTGGTACGCAAAGTCTCTGGAG 603
Qy   130 AlaValTyTrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeu 149
Db   604 GCCGTGGCCATTATGGACTCACTCAAGATCGTCACAGGATCTCAAGCTGGAGAACCTG 663
Qy   150 LeuTyTrLeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLys 169
Db   664 GTTTACTACAACCGCTGGAAGAATCGAAGATTTGTTCATCAGTACGTACTTCCATCTGGCTAAG 723
Qy   170 MetGluGlnAsnGlyIleMetSerThrAlaCysGlyThrProGlyTyTrValAlaProGlu 189
Db   724 CTAGAAA---AATGGCTCATCAAGAGCCCTGTGGACCCCACGAGTACTCTGCCCCAGAG 780
Qy   190 ValLeuLaGlnLysProTyTrSerLysAlaValAspCysTrpSPRIleglyValIlethr 209

```

```
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 47
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-47

Alignment Scores:
Pred. No.: 1-82e-55 Length: 467
Score: 667.00 Matches: 136
Percent Similarity: 93.15% Conservativeness: 0
Best Local Similarity: 93.15% Mismatches: 9
Query Match: 26.54% Indels: 1
DB: 9 Gaps: 0

US-09-960-643-2 (1-476) x US-09-935-464-47 (1-467)

Qy 1 MetGlyArgLysGluGluAspCysSerTrpLysGlnThrThrAsnIleArg 20
Db 30 ATGGTCGAAGGAAGAGATGACTGCAGTTCCTGGGAAGAAACAGACCACCATCCGG 89
Qy 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 90 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCANAAGCTTTTCTCGTG 149
Qy 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
Db 150 AAGCAAGACTGACTGGGAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCACCTGCCTTC 209
Qy 61 ArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
Db 210 CGGACAGCAGCTGNAGATGAGATGCTGTGTTGAAAAGATCAAGCATGAAACATV 269
Qy 81 ValThrLeuGluAspIleTyrgLysThrThrHisTyrrLeuValMetGlnLeuVal 100
Db 270 GTGACCTGGAGGACATCTATGAGAGCACCACCTACTACTGNTCATGCACTGTT 329
Qy 101 SerGlyGlyLeuPheAsp-ArgIleLeuGluArgGlyValTyrrLeuLysAspAl 120
Db 330 TCTGGTGGGAGCTCTTTGACCGGATCTCTGGAGCGGNGTGTCTACACAGAAGATGC 389
Qy 120 aSerLeuValIleGlnValLeuSerAlaValLysTyrrLeuHisGluAsnGlyIleVa 140
Db 390 CAGNCTGGGTGATCCACANGTCTTGTGNCAGTGAATACCTACATGAGATGGCATCGT 449
Qy 140 LHisArgAspLeuLys 145
Db 450 NCACAGAGACTTAAAG 465

RESULT 13
US-09-935-464-48
; Sequence 48, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 48
; LENGTH: 470
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-48

Alignment Scores:
Pred. No.: 2-38e-53 Length: 470
Score: 645.00 Matches: 139
Percent Similarity: 93.29% Conservativeness: 0
Best Local Similarity: 93.29% Mismatches: 6
Query Match: 25.67% Indels: 4
DB: 9 Gaps: 0

US-09-960-643-2 (1-476) x US-09-935-464-48 (1-470)

Qy 1 MetGlyArgLysGluGluAspCysSerTrpLysLysGlnThrThrAsnIle-Ar 20
Db 30 ATGGTCGAAGGAAGAGATGACTGCAGTTCCTGGGAAGAAACAGACCACCATCCCG 89
Qy 20 gLysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGlu-ValPhe-Leu 39
Db 90 GAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAAAGTTTTCCTG 149
Qy 40 VallysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAla 59
Db 150 GTGAAGCAAGACTGACTGGGAGCTCTTGTCTGAGTGCATCAAGAAGTCACCTGCC 209
Qy 60 PheArgAspSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsn 79
Db 210 TTCGGGACAGCAGCTGGGAGATGAGATGCTGTGTTGAAAAGATCAAGCATGAAAC 269
Qy 80 IleValThrLeuGluAspIleTyrrGluSerThrThrHisTyrrLeuValMetGlnLeu 99
Db 270 ATTGTGACCTGGAGGACATCTATGAGAGCACCACCTACTACTCTGGTGCATGCAGCTT 329
Qy 100 ValSerGlyGlyGluLeuPheAsp-ArgIleLeuGluArgGlyValTyrrThrGluLysAs 119
Db 330 GTTCTGTGGGGAGCTCTTTGACCGGATCTCTGGAGCGGNGTGTCTACACAGAAGA 389
Qy 119 pAlaSerLeuValIleGlnValLeuSerAlaValLysTyrrLeuHisGluAsnGlyI 139
Db 390 TGCCAGNCTGGGTGATCCACANGTCTTGTGNCAGTGAATACCTACATGAGATGGCAT 449
Qy 139 eValHisArgAspLeuLys 145
Db 450 CGTNCACAGAGACTTAAAG 468

RESULT 14
US-09-934-406-1
; Sequence 1, Application US/09934406
; Publication No. US20020192204A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 15985, A NOVEL HUMAN SERINE/THREONINE
; FILE REFERENCE: 10448-078001
; CURRENT APPLICATION NUMBER: US/09/934,406
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/226,740
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3552
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (208)....(2505)
US-09-934-406-1

Alignment Scores:
```



```
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7777
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (19)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7777

Alignment Scores:
Pred. No.:      3 07e-52      Length:      474
Score:          633.50      Matches:      118
Percent Similarity: 87.90%      Conservative: 20
Best Local Similarity: 75.16%      Mismatches: 18
Query Match:      25.21%      Indels:      1
DB:               9          Gaps:      1

US-09-960-643-2 (1-476) x US-09-796-692-7777 (1-474)
QY 115 TyrThrGluLysAspAlaSerLeuValIleGlnValLeuSerAlaValLysTyrLeu 134
      |||||.....:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 472 TACACGGAGCGGAGCGCGCGCTCATCTCCAGGTGCTGGATGCTGGAATACCTG 413
QY 135 HisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrPro 154
      |||||.....:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 412 CATGACCTGGCGATGTACACGGGATCTCAAGCCAGAGAACTCTGCTGTACTACGCCCTG 353
QY 155 GluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGly 174
      |||||.....:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 352 GATGAGACTCCAAATCATGATCTCCGACTTGGCCCTCCCAAGATGGAGGCCGGC 293
QY 175 ----IleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGln 193
      |||||.....:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 292 AGTGTGCTCTCACCGCTGTGGAATCCGGGATACGTGGCCCTGAAGTCTCTGGCCAG 233
QY 194 LysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeu 213
      |||||.....:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 232 AAGCCCTACAGCAGGCTGTGGATGCTGTGTCATAGTGTGTCATCGCTACATCTTGCTC 173
QY 214 CysGlyTyrProPheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGlu 233
      |||||.....:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172 TCGGTTACCTCCCTTCTATGACGAGAAATGATGCCAACTCTTTGAACAGATTTCAG 113
QY 234 GlyTyrTyrGluPheGluSerProPheTrpAspPleSerGluSerAlaLysAspPhe 253
      |||||.....:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 112 GCCGAGTACGAGTTTGTACTCTCTACTGGGACGACATCTCTGACTCTGCCAAAGATTTC 53
QY 254 IleCysHisLeuGluLysAspProAsnGluArgTyrThrCysGluLys 270
      |||||.....:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 52 ATCCGGCACTTGATGGAGAGAGACCCAGAAANNNGANTTACCNGTGAGCAG 2
```

Search completed: March 15, 2003, 07:54:16
Job time : 161 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 15, 2003, 05:02:42 ; Search time 64 Seconds
(without alignments)
2280.908 Million cell updates/sec

Title: us-09-960-643-2

Perfect score: 2513

Sequence: 1 MGRKEEDCSSWKQTNR.....VKAGSSHCRAQTGVCLIM 476

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09960643/runat_07032003_090517_19900/app_query.fasta_1.647
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09960643 -CGN1_1_40=runat_07032003_090517_19900 -NCPU=6 -ICPU=3
-NO_MLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	39.4	1282	2	US-08-878-989-12
2	990	39.4	1282	4	US-09-272-796-12
3	747.5	29.7	3124	4	US-09-734-030-1
4	664.5	26.4	3471	2	US-08-715-568A-2
5	586.5	23.3	1858	4	US-09-529-093A-1
6	578.5	23.0	1735	4	US-09-746-694-3
7	574.5	22.9	2061	4	US-09-800-960-1
8	562.5	22.4	2934	1	US-08-198-446B-16
9	562.5	22.4	2934	2	US-08-870-693-16
10	542.5	21.6	1333	4	US-09-142-551A-1
11	542.5	21.6	2132	2	US-09-159-385-3
12	542.5	21.6	2132	4	US-09-186-277-3

13	537.5	21.4	1400	1	US-08-464-164-1
14	537.5	21.4	1400	1	US-08-338-057-1
15	537.5	21.4	1400	2	US-08-668-416-1
16	531	21.1	2514	3	US-08-655-352-1
17	531	21.1	2514	4	US-09-258-016-1
18	531	21.1	2514	4	US-09-257-825B-1
19	525.5	20.9	4935	2	US-08-631-097-3
20	525.5	20.9	5886	4	US-08-810-712-9
21	522.5	20.8	1429	2	US-09-159-385-4
22	522.5	20.8	1429	4	US-09-186-277-4
23	510	20.3	1776	3	US-08-655-352-10
24	510	20.3	1776	4	US-09-258-016-10
25	510	20.3	1776	4	US-09-257-825B-10
26	495.5	19.7	2374	4	US-09-347-801-3
27	477	19.0	1929	4	US-09-359-161-4
28	460.5	18.3	1742	4	US-08-557-006C-38
29	460.5	18.3	2652	4	US-08-557-006C-39
30	460.5	18.3	2761	4	US-08-557-006C-24
31	459.5	18.3	1736	4	US-08-557-006C-37
32	459.5	18.3	1783	4	US-08-557-006C-36
33	452	18.0	362	4	US-08-905-223-89
34	451.5	18.0	2549	4	US-09-467-082-3
35	449.5	17.9	1417	1	US-08-713-828-2
36	449.5	17.9	1417	2	US-08-919-627-2
37	449.5	17.9	1417	2	US-09-096-245-2
38	444	17.7	2637	4	US-09-735-934A-1
39	443	17.6	1647	3	US-09-101-146-44
40	441.5	17.6	1349	1	US-07-951-715A-20
41	441.5	17.6	1349	2	US-08-459-448A-20
42	441.5	17.6	1349	3	US-08-459-595A-20
43	441.5	17.6	1349	3	US-08-459-504B-20
44	441.5	17.6	1349	3	US-08-459-444-20
45	441.5	17.6	1349	4	US-09-547-422-20

ALIGNMENTS

RESULT 1
US-08-878-989-12
; Sequence 12, Application US/08878989
; Patent No. 5885803
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,989
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING AGENT:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J

```
;
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT06
; CLONE: 827431
US-08-878-989-12

Alignment Scores:
Pred. No.: 1,29e-86 Length: 1282
Score: 990.00 Matches: 193
Percent Similarity: 74.92% Conservative: 49
Best Local Similarity: 59.75% Mismatches: 73
Query Match: 39.40% Indels: 8
DB: 2 Gaps: 2

US-09-960-643-2 (1-476) x US-08-878-989-12 (1-1282)
QY 13 LysLysGlnThrThrAsnIleArgLysThrPheIlePheMetGluValLeuLysSerGly 32
DB 213 AAGAAACACAGAGGAGCATCAGCAGCGCTTACGAGATCCCGAGAGCGCTCGGCTGGGT 272
QY 33 AlaPheSerGluValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeuLys 52
DB 273 GCCTTCTCGAGGTGGTGTGCCCGAGGAGCGGGCTCCGACACCTCGTGGCCCTCAAG 332
QY 53 CysIle---LysLysSerProAlaPheArgAspSerLeuGluAsnGlnIleAlaVal 71
DB 333 TGCAATCCCAAGAGGCCCTCCGGGGCAAGGAGGCCCTGTGGAGAACGAGATCGCAGTG 392
QY 72 LeuLysLysIleLysGluAsnIleValThrLeuGluAspIleTyrGluSerThrThr 91
DB 393 CTCCTAGGATCAGTCACCCCAACATCGTCGCTCGGAGGATGCCACGAGAGCCCTTC 452
QY 92 HisTyrTyrLeuValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIleLeuGlu 111
DB 453 CACCTCTACCTGGCCATGGAACCTGTCAGGGTGGCGAGCTGTTTGACCGCATCATGG 512
QY 112 ArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnValLeuSerAlaVal 131
DB 513 CCGGCTCTCTACAGAGAAGGATGCCAGCATCTGGTGGTCAAGGTCTTGGCGCGTC 572
QY 132 LysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLysTyr 151
DB 573 TCCTACCTGCACAGCGCTGGGATCGTCACCGGACCTCAAGCCCGAAACCTCCTGTAT 632
QY 152 LeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeuSerLysMetGlu 171
DB 633 GCCAGCCCTTTGAGGATCGAAGATCATGCTCTGACCTTGGACTCTCAAAATCCAG 692
QY 172 GlnAsnGlyIleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValLeu 191
DB 693 GCTGGAACATGCTAGGACCGCCCTGTGGGACCCCTGGATATGTGGCCCGAGAGCTCTG 752
QY 192 AlaGlnLysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThrTyrIle 211
DB 753 GAGCAGAAACCTTACGGAAGCCGATAGATGTGGGCCCTTGGCGGTCATCTCCTACATC 812
QY 212 LeuLeuCysGlyTyrProPheTyrGluGluThrGluSerLysLeuPheGluLysIle 231
DB 813 CTGCTGTGGGTACCCCTTCTACGACGAGAGCGACCTGAGCTCTTCAGCCAGATC 872
QY 232 LysGluGlyTyrTyrGluPheGluSerProPheTrpAspIleSerGluSerAlaLys 251
DB 913 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
```

```
DB 873 CTGAGGGCCAGCTATGAGTTGACTNCTCTTCTGGGATGACATCTCAGATCAGGCAA 932
QY 252 AspPheIleCysHisLeuLeuLysAspProAsnGluArgTyrThrCysGluLysAla 271
DB 933 GACTTTATTTGGCACCTTCTGGAGCGAGACCTTCAGAAGAGGTTCACTGCCAACAGGCC 992
QY 272 LeuSerHisProTyrPheAspGlyAsnThrAlaLeuHisArgAspIleTyrProSerVal 291
DB 993 TTGCGGAGACCTTTGGATCTTTTGGGACACAGGCTTTGGCAGGACATCTTAGGGTTTGT 1052
QY 292 SerLeuGlnIleGlnLysAsnPheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAla 311
DB 1053 ACTGAGCAGATCCGAGACACTTTGCTTGGACACACTGGAAGCGAGCTTCAATGCCACC 1112
QY 312 AlaValValHisHisMetArgLysLeuHisMetAsnLeuHisSerProGlyValArgPro 331
DB 1113 TTGTTCTCGCCACATCCGGAAGCTG-----GGCAGATCCCA 1151
QY 332 GluValGlu 334
DB 1152 GAGGGCGAG 1160
RESULT 2
US-09-272-796-12
; Sequence 12, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT06
; CLONE: 827431
```



```
Db 664 GTTTACTACAACGGCTGAAGAACTCGAAGATTGTCATCAGTCACTTCCATCTGGCTAAG 723
Qy 170 MetGluGlnAsnGlyIleMetSerThraAlaCysGlyThrProGlyTyrValAlaProGlu 189
Db 724 CTAGAA---AATGGCTTCTCAAGAGCCCTGTGGAGCCCGGAGTATCTGGCCCGAG 780
Qy 190 ValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThr 209
Db 781 GTGGTAGCGCGGACGGGTATGGAGCCCTGTGGACTGCTGGCCATTGGAGTCATCATG 840
Qy 210 TyrIleLeuLeuCysGlyTyrProProPheTyrGluGluThrGlu----- 224
Db 841 TACATCTCTGCTTTCAGGCAATCCACCTTCTATGAGGAGGTGGAAGAGATGATTATGAG 900
Qy 225 -----SerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSerPro 241
Db 901 AACCATGATAGATCTCTCCGAAGATCCTGGCTGGTGTGACTATGATTTTGACTCTCCA 960
Qy 242 PheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLysLeuLeuLysAsp 261
Db 961 TATTGGATGATATTTCAGCGAGCCCAAGACCTGTGTCACAAGGCTGATGGAGGTGGAG 1020
Qy 262 ProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsnThr 281
Db 1021 CAAGACCAAGCGGATCATTGCAAGAGAGGCCATCTCCCATGATGGATTTCTGGCAATGCT 1080
Qy 282 AlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLys 301
Db 1081 GCTTCTGATAGAACATCAAGATGGTGTCTGTGCCAGATGAAAGAACTTTGCCAGG 1140
Qy 302 SerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArg----- 318
Db 1141 GCCAAGTGGAAAGAGGCTGTCGAGTAGCACCTCATGAACAGGCTCCGGGCACACAGAG 1200
Qy 319 -----LysLeuHisMetAsnLeuHisSerProGly 328
Db 1201 CAGTCCAGCAGCGGTGCAGCCCGAGTGCCTCAGC-CACAGACACTGCCACCC-CCGGGG 1258
Qy 329 -----ValArgProGluValGluAsnArg----- 336
Db 1259 CTGCAGGTGGGCCACAGCTGCAGTGGAGTGAGTGAGTACCTCAGCCCTCAGGGTGATG 1318
Qy 337 -----ProProGluThrGlnAlaSerGluThrSer 346
Db 1319 CTGCTCGTCTGCAAGAGTGATATGTGGCCCGCCGACAGCCGTAGTG----- 1366
Qy 347 ArgProSerSerProGluIleThrIleThrGluAlaProValLeuAspHisSerValAla 366
Db 1367 ---CCACCCAGCCACAGATGAAGTGCACCCAGCCACTGTATGGCAGTGTCAACCCAG 1423
Qy 367 LeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArgProThrAlaProGlyGly 386
Db 1424 CCACCGATGGAAGCATCACTCCAGCCACTGATGGAGGTGTCAACCCAGTCACTACAGGA 1483
Qy 387 ArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSerSer-----SerLeuVal 404
Db 1484 GCGCTACTCCAGCCACTATG---GGAGAGCCACACCA-GCCACAGAAGAGAGACTGTG 1539
Qy 405 ProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSerSerCysLeuAsn 424
Db 1540 CCCACCCCAAGCAGTGCCAT-GCTGGCCACCAAGGCGAGTGCACCCCTGAGCCCGG 1598
Qy 425 IleGlySerLysGlyLysSerSer 432
Db 1599 TATGGCCCGAGCGGACAGCACAGC 1622
```

RESULT 4

```
US-08-715-568A-2
; Sequence 2, Application US/08715568A
; Patent No. 5856463
; GENERAL INFORMATION:
; APPLICANT: Prydz, Hans Peter Blankenborg
```

```
; APPLICANT: Brede, Gaute
; TITLE OF INVENTION: PSKH-1 Ribozymes and Uses in Disease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik
; STREET: 600 South Avenue West
; CITY: Westfield
; STATE: NJ
; COUNTRY: USA
; ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/715,568A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Foley, Shawn P.
; REFERENCE/DOCKET NUMBER: FORSK 3.0-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-654-5000
; TELEFAX: 908-654-7866
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3471 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-715-568A-2
```

```
Alignment Scores:
Pred. No.: 2 33e-54 Length: 3471
Score: 664.50 Matches: 150
Percent Similarity: 54.95% Conservative: 72
Best Local Similarity: 37.13% Mismatches: 123
Query Match: 26.44% Indels: 59
DB: 2 Gaps: 9
```

```
US-09-960-643-2 (1-476) x US-08-715-568A-2 (1-3471)
Qy 27 GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly 46
Db 398 GAAATAATTGGCGGAGGAGCGAGCTTCAGCCGAGTGTGTAGTGTAGAGCACCGGCAACCCGG 457
Qy 47 LysLeuPheAlaLeuLysCysIleLysLysSerProAlaPheArgAspSerSerLeuGlu 66
Db 458 CAGCCGTATGCCATCAGATGATTGAGACCAAGTACCGGGAGGGCGGGAGGTGTGTGAG 517
Qy 67 AsnGluIleAlaValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIle 86
Db 518 TCGGAGCTGCGTGTGCTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTG 577
Qy 87 TyrGluSerThrThrHisTyrTyrLeuValMetClnLeuValSerGlyGlyGluLeuPhe 106
Db 578 TTCGAGACACAGGAGCGGGGTGTACATGGTGTGAGCTGGCCACTGGTGTGAGAGCTCTTT 637
Qy 107 AspArgIleLeuGluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnGln 126
Db 638 GACCGCATCATGCCAAGGGCTCTCTTCCAGGAGGTGACCGCCAGCGGGGTGCTGCAGATG 697
Qy 127 ValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysPro 146
Db 698 GTGCTGGATGGCGTCCGGTATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 757
Qy 147 GluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIleMetIleThrAspPheGly 166
Db 758 GAGAAATCTGCTTACTACCATCCGGGCACTGACTCCAAAGATCATCATCCGCTTCCGGC 817
Qy 167 LeuSerLysMetGluGlnAsnGly-----IleMetSerThrAlaCysGlyThrPro 183
```

```

|||||
Db 818 CTGGCCAGTCTCCCAAGAGGGTGATGACTGCTGTGATGAAGACCACCTGTGGCAGCCCT 877
QY 184 GlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaValAspCysTrp 203
Db 878 GAGTACATTGCCCAAGAGTCTGGTCCGAAGCCATACCAACTCAGTGGACATGTGG 937
QY 204 SerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyrGluGluThr 223
Db 938 CGCTGGCGGTCATTCGCTACATCTACTCAGTGGCCACCACCGCTTTGAGGATGACAAC 997
QY 224 GluSerLysLeuPheGluGlyLysIleLysGluGlyTyrTyrGluPheGluSerProPheTrp 243
Db 998 CGTACCGGCTGTACCGGCAGATCTCTCAGGGCAAGTACAGTTACTCTGGGGAGCCCTGG 1057
QY 244 AspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLysAspProAsn 263
Db 1058 CCTAGTGTGCTCAACCTGGCCAGGAGCTTCATTGACCGCTCTGACAGTGGACCCCTGGA 1117
QY 264 GluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsnThrAlaLeu 283
Db 1118 GCCGTGATGACTGCTGAGGCCCTGAGGCACCCGGGGTG----- 1159
QY 284 HisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAlaLysSerLys 303
Db 1160 -----GTGAGCATG----- 1168
QY 304 TrpArgGlnAlaPheAsnAlaAlaValValHisMetArgLysLeuHisMetAsn 323
Db 1169 -----GTCGCTCTTCATCCATGAAGAACCCTGCACCGCTCC 1204
QY 324 LeuHisSerProGlyValArgProGluValGluAsnArgProProGluThrGlnAlaSer 343
Db 1205 ATA---TCCAGAACCTCTTAACGTGCTCTCGGCTGCCAGACCAACCTGCC 1261
QY 344 GluThrSerArg-ProSerSerProGluIleThrIleThrGluAlaProValLeuAspH1 363
Db 1262 CAGTCCACGCGTCCAGCCGCTCCA-----CAGCGCTCCAATA---AGTCA 1303
QY 363 sSerVal-AlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArgProThra 383
Db 1304 CGCCGTGCGGGGAACGGAGCTCG-----GGAGCTCAACCTCGCG 1345
QY 383 laProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSerSerL 403
Db 1346 TACCAGCA-----GCAATCAATGCTGAGCCGAATGGCTGTGCACATGCGACGCA 1399
QY 403 euValProMethHis-----GlnGlySerLeuAlaAlaGlyP 415
Db 1400 CCCAGCCTGGCCACACACTGTGTGTCATCTGGTCCGATCGCTCTCTGGAGATAGGCG 1459
QY 415 roCysGly 417
Db 1460 TATGTGGC 1467

RESULT 5
US-09-529-093A-1
; Sequence 1, Application US/09529093A
; Patent No. 6413755
; GENERAL INFORMATION:
; APPLICANT: LUYTEN, Walter H.M.L.
; APPLICANT: PARKER, Andrew E.
; APPLICANT: MCGOWAN, Clare H.
; APPLICANT: BLASINA, Alessandra
; TITLE OF INVENTION: Human Checkpoint Kinase, hCDS1, Compositions and Methods
; FILE REFERENCE: TSRI 649.0
; CURRENT APPLICATION NUMBER: US/09/529,093A
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: PCT/EP98/06981
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: GB 9722320.0
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 6
;
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (66)..(1694)
US-09-529-093A-1
Alignment Scores:
Pred. No.: 3,17e-47 Length: 1858
Score: 586.50 Matches: 142
Percent Similarity: 52.60% Conservative: 60
Best local Similarity: 36.98% Mismatches: 131
Query Match: 23.34% Indels: 51
DB: 4 Gaps: 11
US-09-960-643-2 (1-476) x US-09-529-093A-1 (1-1858)
QY 7 AspAspCysSerSerTrpLysLysGlnThrThrAsnIleArgLysThrPheIlePheMet 26
Db 684 GATGATCAGTCAGTTTATCTTAAGGCA-----TTTAAGAGATGAATACATCATGTCA 734
QY 27 GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly 46
Db 735 AAAACTCTTGGAGTGGTGGCTGTGGAGAGGTAAAGCTTTCGAGAGGAGAAACATGT 794
QY 47 LysLeuPheAlaLeuLysCysIleLysLys----- 56
Db 795 AAGAAGTAGCCATAAAGATCATCATCAGCAAAAGGAAGTTTGCATTATGGTTACGAAGAGAG 854
QY 57 ---SerProAlaPheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIle 75
Db 855 CGAGACCAGCTCTC-----AATGTTGAAACAGAAATAGAAATTTTGAAGAACGTA 905
QY 76 LysHisGluAsnIleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeu 95
Db 906 AATCATCTCTGCATCATCAAGATTAACACTTTTGTGATGCA---GAAGATTATTATTATT 962
QY 96 ValMetGlnLeuValSerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyr 115
Db 963 GTTTTGGAAATTGATGAAAGGGGAGAGCTGTTTGAACAAAGTGGTGGGAATAAACGCGCTG 1022
QY 116 ThrGluLysAspAlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHis 135
Db 1023 AAAGAAGTACCTCGCAAGCTCTATTTTACCAGATGCTCTTGGCTGTGCAGTACCTTCAT 1082
QY 136 GluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGlu 155
Db 1083 GAAACGGTATTATACACCGTGACTTAAAGCCAGAGAAATGTTTACTGTCTATCTCAAGAA 1142
QY 156 GluAsnSerLysIleMetIleThrAspPheGlyLeuSerLys---MetGluGlnAsnGly 174
Db 1143 GAGGACTGCTTATAAAGATTACTGATTTTGGCAGCTCCAAGATTTTGGGAGAGACCTCT 1202
QY 175 IleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGln--- 193
Db 1203 CTCATGAGAACCCTTATGTGGAACCCCCACCTACTTGGCGCTGAGATTTGTTTCTGTGTT 1262
QY 194 -----LysProTyrSerLysAlaValAspCysTrpSerIleGlyValIleThrTyrIle 211
Db 1263 GGGACTGTGGGTATAACCGTGCTGTGGACTGTGGAGTTTGAAGGATTATTCTTTTATC 1322
QY 212 LeuLeuCysGlyTyrProProPheTyrGlu---GluThrGluSerLysLeuPheGluLys 230
Db 1323 TGCCTTAGTGGGTATCCACCTTCTCTGAGCATAGGACTCAAGTGTCACTGAAGGATCAG 1382
QY 231 IleLysGluGlyTyrTyrGluPheGluSerProPheTrpAspPheIleSerGluSerAla 250
Db 1383 ATCCACCATGGAAATAACAACCTTCATTCCTCCTGAGTCTGGCGAGAGTCTCAGAGAAAGCT 1442
QY 251 LysAspPheIleCysHisLeuLeuLysAspProAsnGluArgTyrThrCysGluLys 270
```


QY 17 ThrAsnIleArgLysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGlu 36
DB 134 ACCGGTTTCACCGACGACATACAGCTCTTCGAGGAGCTTGGCAAGGGTCTTCTCTGTG 193
QY 37 ValPheLeuValLysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIle----- 54
DB 194 GTCGCGAGGTGTGTGAAGAAACCTCCACGCGAGGAGTAGCGCAAAATCATCAATACC 253
QY 55 LysLysSerProAlaPheArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLys 74
DB 254 AAGAAATTGCTGCCCGGATACACAGAACTAGACGTGAGGCTCGGNATATGTCGACTT 313
QY 75 IleLysHisGluAsnIleValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyr 94
DB 314 CTGAACATCCAAACATCGTGGCCCTCCATGACAGATATTCTGAAGAGGGTTTCACATAC 373
QY 95 LeuValMetGlnLeuValSerGlyGluLeuPheAspArgIleLeuGluArgGlyVal 114
DB 374 CTCGTGTTTGACCTGTGTACCGCGGGAGCTGTTTGAAGACATTTGTGCCACAGAGTAC 433
QY 115 TyrThrGluLysAspAlaSerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeu 134
DB 434 TACAGTGAACGAGATGCCAGCCACTGTATACATCAGATTCTGGAGAGTGTAAACCAATC 493
QY 135 HisGluAsnGlyIleValHisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrPro 154
DB 494 CACCAGCATGACATCGTCCACAGGAGCTGAAGCTTGAGACCTGCTGTGGCGAGTAA 553
QY 155 GluGluAsnSerLysIleMetIleThrAspPheGlyLeuSer-----LysMetGlu 171
DB 554 TGAAGGCTGCCCGCTGAAGCTGGCTGATTTGGCTAGCCATCGAAGTACAGGAGAG 613
QY 172 GlnAsnGlyIleMetSerThrAlaCysGlyThrProGlyTyrValAlaProGluValLeu 191
DB 614 CAGCAGCTGTTGGTGTGTTTGTCT--GGCACCCAGGTACTGTGCCCTGAGGTCTTG 670
QY 192 AlaGlnLysProTyrSerLysAlaValAspCysTyrSerIleGlyValIleThrTyrIle 211
DB 671 AGAAAGATCCCTATGAAAACCTGTGGATATCTGGGCTGGGGTCTATCTGTATATC 730
QY 212 LeuLeuCysGlyTyrProPheTyrGluThrGluSerLysLeuPheGluLysIle 231
DB 731 CTCCTGTGGGCTATCTCTCTCTGGATGAGGATCAGCACAGCTGTATCAGCAGATC 790
QY 232 LysGluGlyTyrTyrGluPheGluSerProPheTyrPaspAspIleSerGluSerAlaLys 251
DB 791 AAGCTGGAGCTATGATTTCCATCACCAGATGGGACACGGTAACTCCTGAAGCCAG 850
QY 252 AspPheIleCysHisLeuLeuGluLysAspProAsnGluArgTyrThrCysGluLysAla 271
DB 851 AACTGTATCAACAGATGCTGACCAATAAACCCAGCAAGCGCATCAGGCTGACCAAGGT 910
QY 272 LeuSerHisProTyrIleAspGlyAsnThrAlaLeuHisArgAspIleTyrProSerVal 291
DB 911 CTCAAGCACCGGTGGGTCTGTCAACGATCCACGCTGGCATCCATGATGCTGTCAGGAG 970
QY 292 SerLeuGlnIleGlnLysAsnPhe--AlaLysSerLysTyrArgGlnAlaPheAsnAla 310
DB 971 ACTGTGGAGGTGTTGGCCAGATGCAATGCCCGGAGAAACTGAAGGGTGCCTATCTCAG 1030
QY 311 AlaAlaValValHisHisMetArgLysLeuHisMetAsnLeuHisSerProGlyValArg 330
DB 1031 ACCATGCTGTCTCCAGAACTTCTCAGTTGGCAGGACGAGCTCCGCCCC--GCCTCG 1087
QY 331 ProGluValGluAsnArgProProGluThrGlnAlaSerGluThrSer----- 346
DB 1088 CTGCCGCG-GAGCGCCCGCGGCTGCGCGGCAAGCTGCGCAAAAGCCTATTGAACAAGAA 1146
QY 347 -----ArgProSerProGluIleThrIleThrGluAlaProValLeuAspHis 363
DB 1147 GTCGATGGCGGTGTCAGAAAGGAAGTTCAGTTCCAGGTGTCACCTAA-TGGAGCCAC 1205

QY 364 SerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArgProThrAla 383
DB 1206 AAACCACTGTGGTACACACAGCTACAGATGGGATCAAGGGCTCCACAGAGCTGCAACA 1265
QY 384 ProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSerSerLeu 403
DB 1266 CCACCAAGAAG----- 1277
QY 404 ValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSerSerCysLeu 423
DB 1278 -----ATGAGGACCTCAAGCTGCCGCTCGCACCTGGGAATGGCAGCTCGGTGCTG 1331
QY 424 AsnIleGlySerLysGlyLysSer-----SerTyrCysSerGluProThrLeu 439
DB 1332 AAGCAGCAGGCTCCGCGCAGACAGACAGCCCTCTCAGCATGTCAGC---CCAGCCTT 1388
QY 440 LeuLysLysAlaAsnLysLysGlnAsnPheLysSer 451
DB 1389 CTCTCTCTCTCAGCCATCGCAAAACAGGAGATCA 1424
RESULT 8
US-08-198-446B-16
; Sequence 16, Application US/08198446B
; Patent No. 5674996
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Pilon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/198,446B
; FILING DATE: 18-FEB-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FHCRI7537
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
; TELEFAX: 206-224-0779
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2934 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; DESCRIPTION: yeast MEC2 cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 395..2724
US-08-198-446B-16
Alignment Scores: 1.39e-44 Length: 2934
Pred. No.:

Score: 562.50 Matches: 157
Percent Similarity: 48.50% Conservative: 70
Best Local Similarity: 33.55% Mismatches: 165
Query Match: 22.38% Indels: 78
DB: 1 Gaps: 15

US-09-960-643-2 (1-476) x US-08-198-446B-16 (1-2934)

```
QY 10 SerSerTriPlysLysGlnThrThrAsnIleArgLysThrPhe---IlePheMetGluVal 28
Db 944 TCATCAATGGTGCACCAAGACTGTGATTTTAAGGATTTTCGATATTATGACGAAGTG 1003
QY 29 LeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGlyLysLeu 48
Db 1004 GTGGCCAGGGTGCATTTGCCAGTAAAGAAAGCCATTGAAAGAACTACTGGGAAACA 1063
QY 49 PheAlaLeuLysCysIleLysLysSerProAlaPheArgAsp---SerSerLeuGluAsn 67
Db 1064 TTCGCGGTGAAGATATATAAGTAACCAAGTAATAGGCAATATGATGGTGTGACAGA 1123
QY 68 GluIleAlaValLeuLysLysIleLysHisGluAsnIleValThrLeuGluAspIleTyr 87
Db 1124 GAGTTAGAAGTATGCAAAAGCTCAATCATCAAGGATAGTACGATTGAAAGGATTTTAT 1183
QY 88 GluSerThrThrHisTyrTyrLeuValMetGlnLeuValSerGlyCysLeuPheAsp 107
Db 1184 GAAGATACTAGAGATATTATATGTTGGTGGAGTTCTGTTCTGGTGGTCAATGAT 1243
QY 108 ArgIleLeuGluArgGlyValTyrThrGluLysAspAlaSerLeuValIleGlnVal 127
Db 1244 TTGTTGCTGCTCATGGTGGGTGGAGAGATGCTGGAGGAGATATCCAGGAGATA 1303
QY 128 LeuSerAlaValTyrTyrLeuHisGluAsnGlyIleValHisArgAspLeuLysProGlu 147
Db 1304 CTCACAGCAATAAATACATTCATCTATGGGCATCAGCCATCGTACCTAAAGCCGAT 1363
QY 148 AsnLeuLeuTyrLeuThrProGluGluAsnSerLysIleMetIleThrAspPheGlyLeu 167
Db 1364 AATATCTCT---AATGAACAAGCATCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1420
QY 168 SerLysMetGluGlnAsnGly---IleMetSerThrAlaCysGlyThrProGlyTyrVal 186
Db 1421 GCAAAAGTACAAGAAATGGTCTTTTATGAAACCTTCTGTGGCATTGTCATATGTG 1480
QY 187 AlaProGluValLeuAlaGlnLys----- 194
Db 1481 GCACCTGAAGTCATCAGAGTAAGATACATCCGTATCTCTGATGAATACGAAGAAAG 1540
QY 195 ---ProTyrSerLysAlaValAspCysTyrSerIleGlyValIleThrTyrIleLeuLeu 213
Db 1541 AATGAGTACTCTCGTTAGTGGATATGTTGTTCAATGGGATGTTGTTGTTGTTGTTGTT 1600
QY 214 CysGlyTyrProPheTyrGluThrGluSerLysLeuPheGluLysIleLysGlu 233
Db 1601 ACGGGCCACTTACCTTTTAGTGTACACACAGAGCACTATATATAACAGATTTGAAGA 1660
QY 234 GlyTyrTyrGluPheGluSerProPheTyrAsp---AspIleSerGluSerAlaLysAsp 252
Db 1661 GGCTCATATCAT---GAAGGGCCCTCAAGATTTCCGGATATCTGAAGAGCAAGAGAT 1717
QY 253 PheIleCysHisLeuGluLysAspProAsnGluArgTyrThrCysGluLysAlaLeu 272
Db 1718 TTCAATGATTCATGTTACAGGTGGATCCAAATATAGATTCGACAGCTGCAAAAGCCCTG 1777
QY 273 SerHisProTyrIleAspGlyAsnThrAla-LeuHisArg----- 285
Db 1778 AATCATCCCTGGATCAAGATGATGCTCCATCATCATATATGTTGATTTTCACAA 1837
QY 286 -----AspIleTyrProSerValse 292
Db 1838 ATATCCTTTATCAATCGTTGTGGCAGCAGAAATTTATGAATAATATGACGATGCTCAA 1897
QY 292 rLeuGlnIleGlnLysAsnPheAlaLysSerLysTyrArgGlnAlaPheAsnAlaAla 312
```

```
Db 1898 TACGAATTTGTCAAAGCGCAAGAAATTTACAAATGAGCAACAACCTTCAAG----- 1949
QY 312 aValValHisHisMetArgLysLeuHisMet----- 322
Db 1950 -----AACAGGATCAGGAAGCAACCAAGATGGAAATTCAGGATTTAAATACCGCA 2002
QY 323 -----AsnLeuHisSerPro---GlyValArgProGluValGluAsnArg--- 336
Db 2003 CACGCCCTATTTCATATACACAGCCCAAAAGCATTTGAAGCAGAACTAGAGAACAATA 2062
QY 337 -----ProGluThrGlnAlaSerGluThrSerArgProSerSerProGluL 353
Db 2063 CTTTTCATCTCCAT-AATACTGAGATTCGAAGAGCTCAAGAGCTCAAGAAAGGTTAATG 2121
QY 353 eThrIleThrGluAlaProValLeuAspHisSerValLeuProAlaLeuThrGlnLe 373
Db 2122 GTTTTACITTAACCAATGCTGACAGCATTAATTCAA-----GAAAGCCT 2169
QY 373 uProCysGlnHisGlyArgArgProThrAlaProGly---GlyArgSerLeuAsnCysLe 392
Db 2170 GGAGATTCAGCAAGGTGTGAATCCATTTTTCATTTGTTGATGATCCGAGGATTCATTT 2229
QY 392 uValAsnGlySerLeuHisIleSerSerSerSerLeuValProMetHisGlnGlySerLeu 412
Db 2230 AATTGAAGACAAATAGTT-GTCTCGAGTTTCATTCCTTCATTTTCAAAAAGAGGCA----- 2283
QY 412 aAlaGlyProCysGlyCysCys 419
Db 2284 -----TGCTGT 2289

RESULT 9
US-08-870-693-16
; Sequence 16, Application US/08870693
; Patent No. 5866338
; GENERAL INFORMATION:
; APPLICANT: Hartwell, Leland H.
; APPLICANT: Weinert, Ted A.
; APPLICANT: Pilon, Sharon E.
; APPLICANT: Groudine, Mark T.
; TITLE OF INVENTION: Cell Cycle Checkpoint Genes
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/870,693
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/198,446
; FILING DATE: February 18, 1994
; APPLICATION NUMBER: PCT/US93/04458
; FILING DATE: May 12, 1993
; APPLICATION NUMBER: US 07/884,426
; FILING DATE: May 14, 1992
; APPLICATION NUMBER: US 07/882,051
; FILING DATE: May 12, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: FPCR110798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-682-8100
```

Db	Seq ID	Accession	Length	Score	E-value	Identity	Positives	Gaps	Strandedness	Topology	Molecule Type	Description	Organism	Feature	Name/Key	Location	US-08-870-693-16
Db	1601	ACGGCCACCTACCTTTTACGTAGTACACACAGCACCATTATATATAACAGATTGGAAGA	1666			100.00	1666	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	234	GlyTyrTrpGluPheGluSerProPheTrpAsp---AspIleSerGluSerAlaLysAsp	252			100.00	252	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Db	1661	GGCTCATATCAT---GAAGGGCCCTCAAGATTTCCGGATATCTCGAAGCAAGAGAT	1717			100.00	1717	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	253	PheIleCysHisLeuLeuGluLysAspProAsnGluArgTrpThrCysGluLysAlaLeu	272			100.00	272	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Db	1718	TTTCATAGATTTCATTTTACAGGTGGATCCAAATAAATAGGTGCACAGCTCCAAAGCCTTG	1777			100.00	1777	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	273	SerHisProTrpIleAspGlyAsnThrAla-LeuHisArg-----AspIleTrpProSerValse	285			100.00	285	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Db	1778	AATCATCCCTGGATCAAGATGAGTCCATTGGCTGCACAAATCATATGTTGATTTTTCACAA	1837			100.00	1837	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	286	-----AspIleTrpProSerValse	292			100.00	292	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Db	1838	ATATCCTTATCACAATCGTTGTCGACGACGAAATTTATAGAAATATGGACGATGCTCAA	1897			100.00	1897	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	292	rLeuGlnIleGlnLysAsnPheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaL	312			100.00	312	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Db	1898	TACGAATTTGTCAAGCGCAAGAAATTTACAAATGAGGACCAACTTCAAG-----	1949			100.00	1949	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	312	aValValHisHisMetArgLysLeuHisMet-----	322			100.00	322	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Db	1950	-----AACAGGATCAGGAAGCAAGATGGAAAAATTCACGATTTAAAAATACCGCA	2002			100.00	2002	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	323	-----AsnLeuHisSerPro---GlyValArgProGluValGluAsnArg--	336			100.00	336	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Db	2003	CACGCCCTTATTCATATACACGCCCAAGAAAGCATTTGAAGCAGAACTAGAGAACAAAA	2062			100.00	2062	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	337	-----ProProGluThrGlnAlaSerGluThrSerArgProSerProGluI	353			100.00	353	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Db	2063	CTTTTACATTCAT---AATACTGAGATGTCAAGAGCTCAAGAAAGGTAATGGTAG	2121			100.00	2121	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	353	eThrIleThrGluAlaProValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLe	373			100.00	373	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Db	2122	GTTTTAACTTTAAACCATTCGCTGACGATTTTCAA-----GAAAGCCT	2169			100.00	2169	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	373	uProCysGlnHisGlyArgArgProThrAlaProGly---GlyArgSerLeuAsnCysLe	392			100.00	392	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Db	2170	GGAGATTCAGCAAGGTGTGAATCCATTTTCATTTGATGATCCGAGGATTCGAATTGTA	2229			100.00	2229	0	single	linear	cdna	yeast MEC2 cdna	Saccharomyces cerevisiae			395..2724	
Qy	392	uValAsnGlySerLeuHisSerSerSerSerSerSerSerSerSerSerSerSerSerSer	412			100.00	412	0	single	linear							

Db	382	TC	TGCGCGGGAGCTCTTTGACTTCCTGCGGGAGAAAGAGTCGCTGACGGAGACGAGGCC	441
Qy	121	ser	LeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal	140
Db	442	ACC	CAGTTCTCAAGCAGATCTTGGAGGCGTCTACTACCTGCACCTCTAAGCGCATCGCA	501
Qy	141	His	ArgAspLeuLysProGluAsnLeuTyrLeuThrProGlu---GluAsnSerLys	159
Db	502	CAC	TTTGACCTGAAGCCGGAACATCATGTGCTGGACAGAAGCGTCCCAACCCACGA	561
Qy	160	Ile	MetIleThrAspPheGlyLeuSer---LysMetGluGlnAsnGlyIleMetSerThr	178
Db	562	AT	CAAGCTCATCGACTCGGCATCGGCACAGATCGAGCGGGGACGAGTTCAAGAAC	621
Qy	179	Ala	CysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys	198
Db	622	AT	CTTCGGCACCCCGAGTTGTGGCCCCAGAGATGTGAACATFAGACCGCTGGCGTG	681
Qy	199	Ala	ValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro	218
Db	682	GAG	CGGACATGTGGAGCATCGGTGCATCACCATATATCTCTCTGAGCGGTGCATCCCG	741
Qy	219	Phe	TyrGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe	238
Db	742	TT	CTTCGGCGAGACCAAGCAGGAGCGCTTCACCAACATCTCAGCGCTGAAGTACG	801
Qy	239	Glu	SerProPheTrpAspPheSerGluSerAlaLysAspPheIleCysHisLeuLeu	258
Db	802	GAG	CAGGAGTACTTCACACACCGCAGCGAGCTGGCCAGGACTTCATTTCGCCGCTGC	861
Qy	259	Glu	LysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp	278
Db	862	GT	CAAGAATCCCAAGCGGAGATGACCATTCGCCAGAGCCTGGAAACATTCCTGGATTAA	920
Qy	279	Gly	AsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGluIleGlnLysAsn	298
Db	921	GG	CGATCCCGCG-----CGCGATCCCGCG-----CGCGATCCCGCG-----	932
Qy	299	Phe	AlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValAlaHisHisMetArg	318
Db	933	---	CGCGAACGTGCGTG-----CGCGAACGTGCGTG-----CGCGAACGTGCGTG	947
Qy	319	Lys	LeuHisMetAsnLeuHisSerProGlyValArgProGluValGluValAsnArgProPro	338
Db	948	---	-----TGAGCAGCGCGCGCAAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCG	983
Qy	339	Glu	ThrGlnAlaSerGluThrSerArgProSerProGluIleThrIleThrGluAla	358
Db	984	GA	AGACCAACGCGTCTGAAGG-AGTACACCATCAAGT-----CGCACT	1024
Qy	359	Pro	ValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuPro---CysGlnHis	377
Db	1025	CC	AGCTTCGCGCCCAACACAGCTACGCGCTTCGAGCGCTTCCTCCAAAGTGTGGAGG	1084
Qy	378	Gly	ArgArgProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGly-SerLe	397
Db	1085	AG	CGCGCGCGCG-----CCGAGGAGGCGCTCGCGAGCTCGCGCGAGCGCGCGGTCT	1138
Qy	397	u	-----HisIle-----SerSerSerLeuValProMet	406
Db	1139	GC	CACGAGGACGTGGAGGCGCTGGCGGCATCTACGAGGAGAGGAGGCGCTGGTACCG	1196
Qy	406	thi	GlnGlySerLeuAlaAlaGlyProCysGlyCysSerSerCysLeuAsnIleGln	426
Db	1197	---	CGAGGAGCGACACGCTGGCGCAGG-ACCTGC-----GG	1230
Qy	426	ysr	LysGlyLysSerSerTyrCysSerGluProThrLeuLeuLysLysAlaAsnLysLys	446
Db	1231	AG	GTACGCGAGGAGG-----TGCTCAGACCGAGG---CGCTCAACCGCAGCGCGAG	1281
Qy	446	sgl	AsnPheLysSerGluValMetValProValLysAlaSerGlySerSer	463
Db	1282	GAG	AGGCGCAAGGGCGCGTGTGGGGACGCGGCTTCAGCGCGCGCTTCA	1333

RESULT 12
 US-09-186-277-3
 : Sequence 3, Application US/09186277
 : Patent No. 6171841
 : GENERAL INFORMATION:
 : APPLICANT: AKIRA, SHIZUO
 : APPLICANT: KAWAI, TARO
 : TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
 : FILE REFERENCE: 081356/0128
 : CURRENT APPLICATION NUMBER: US/09/186,277
 : CURRENT FILING DATE: 1998-11-05
 : EARLIER APPLICATION NUMBER: JP97/261589
 : EARLIER FILING DATE: 1997-09-26
 : NUMBER OF SEQ ID NOS: 8
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 3
 : LENGTH: 2132
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (94)..(1455)
 US-09-186-277-3

US-09-960-643-2 (1-476) x US-09-186-277-3 (1-2132)

Qy	27	GluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeuThrGly	46
Db	142	GAGAGCTGGGACGCGCCAGTTGGCATGCTCGGAAGTGC CGGCAAGGCGACGGC	201
Qy	47	LysLeuPheAlaLeuLysCysIleLysLys-----SerProAlaPheArgAspSer	63
Db	202	AAGGAGTACGACCAAGTTCATCAAGAAGCGCGCTGTCAACCAGCGCGTGGGGTG	261
Qy	64	Ser-----LeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle	80
Db	262	AGCGCGGAGGAGATCGACGCGGAGGTGAACATCCTCGGAGATCGGCACCCCAACATC	321
Qy	81	ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal	100
Db	322	ATCACCTGCGACGACATCTTCGAGAACACACGCGGTGCTCTCATCTGGAGCTGGTC	381
Qy	101	SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla	120
Db	382	TCTGCGGGAGCTCTTTGACTCTCTCGCGGAGAAAGAGTCGTGACGGAGGACGAGGC	441
Qy	121	SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal	140
Db	442	ACCCAGTTCCTCAAGCAGATCTCTGGAGCGGTTCATACCTGCATCTAAGCGCATCGCA	501
Qy	141	HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGlu---GluAsnSerLys	159
Db	502	CACTTTCACCTGAGCGCGAAACATCATCTCTCTGGCAAGAACAGTGGCCACCCACGA	561
Qy	160	IleMetIleThrAspPheGlyLeuSer---LysMetGluGlnAsnGlyIleMetSerThr	178
Db	562	ATCAAGTTCATCGACTTCGGCATTCGCGCACAGATCGAGCGGGGAACAGATTCAAGAAC	621
Qy	179	AlaCysGlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLys	198
Db	622	ATCTTCGGCACCCCGGAGTTGTGCCCCAGAGATTGTGAACATGAGCGCGTGGGCCGTG	681
Qy	199	AlaValAspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPro	218

Db 682 GAGCGGACATGTGGAGCATCGGTGTATCATACCTATATCTCTGAGCGGTGCATCCCG 741
Qy 219 PheTyrGluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPhe 238
Db 742 TTCCTGGCGAGCAACAGCAGAGCGCTCACACATCTCAGCGTGTAACATCCGACATTC 801
Qy 239 GluSerProPheTrpAspPheSerGluSerAlaLysAspPheIleCysHisLeuLeu 258
Db 802 GACGAGGAGTACTTACGACACACAGCGAGCTGGCCAGGACTTCATTCGCGGCTGCTC 861
Qy 259 GluLysAspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAsp 278
Db 862 GTCAAGATCCCAAGCGGAGATGACCATTTGCCAGAGCTGGAACATTCCTGGATTAA- 920
Qy 279 GlyAsnThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsn 298
Db 921 GCGGATCCGCGC----- 932
Qy 299 PheAlaLysSerLysTrpArgGlnAlaPheAsnAlaAlaValHisMetArg 318
Db 933 ---GCGGACGTCGTGG----- 947
Qy 319 LysLeuHisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProPro 338
Db 948 -----TCAGGACAGCGCGCCCAAGCCGCGCGCGCT 983
Qy 339 GluThrGlnAlaSerGluThrSerArgProSerSerProGluIlePheIleThrGluAla 358
Db 984 GAAGACACCGCTCTGAAG-AGTACACCATCAAGT-----CGCACT 1024
Qy 359 ProValLeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuPro---CysGlnHis 377
Db 1025 CCAGCTTGGCGCCCAACAGCTACGCCGACTTCGAGCGCTTCCAGGTGCTGGAGG 1084
Qy 378 GlyArgArgProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGly-SerLe 397
Db 1085 AGCGCGCGCGC-----CCGAGGAGCGCTGCGGAGCTGAGCGAGCGCGCGCTCT 1138
Qy 397 u-----Hisile-----SerSerSerLeuValProMe 406
Db 1139 GCACAGGAGCGTGGAGCGCTGGCGCCCATCTACGAGGAGGAGGCGCTGTGTACCG-- 1196
Qy 406 tHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSerSerCysLeuAsnIleGl 426
Db 1197 ----CGAGGAGCGAGCGAGCTGGCGCAGG-ACCTGC-----GG 1230
Qy 426 ySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeuLysLysAlaAsnLysLy 446
Db 1231 AGGCTACGCGGAGGAGC-----TGCTCAAGACCGAGG---CGCTCAAGCGGCGGCGCAG 1281
Qy 446 sGlnAsnPheLysSerGluValMetValProValLysAlaSerGlySerSer 463
Db 1282 GAGGAGCCAAAGCGCGCTGCTGGGAGCAGCGCGCTCAAGCGCGCTTCA 1333

RESULT 13
US-08-464-164-1
; Sequence 1, Application US/08464164
; Patent No. 5614195
; GENERAL INFORMATION:
; APPLICANT: Tonley, Fiona M.
; APPLICANT: Dunn, Paul P. J.
; APPLICANT: Bumstead, Janene M.
; APPLICANT: Vermeulen, Arno N.
; TITLE OF INVENTION: Coccidiosis poultry vaccine
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5614195el patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/464,164
APPLICATION NUMBER: US/08/464,164
FILING DATE: June 2, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1400 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: C-terminal
ORIGINAL SOURCE:
ORGANISM: Eimeria maxima
STRAIN: Houghton
DEVELOPMENTAL STAGE: sporozoite
IMMEDIATE SOURCE:
LIBRARY: sporozoite cDNA cloned in Lambda ZAPII
CLONE: Em70-1
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1368
US-08-464-164-1
Alignment Scores:
Pred. No.: 1,14e-42 Length: 1400
Score: 537.50 Matches: 122
Percent Similarity: 57.70% Conservative: 54
Best Local Similarity: 40.00% Mismatches: 116
Query Match: 21.39% Indels: 13
DB: 6

US-09-960-643-2 (1-476) x US-08-464-164-1 (1-1400)
Qy 25 PheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuValLysGlnArgLeu 44
Db 4 TTCGTGGAAGTTTGGGT-----GAGGTCTCTTATGCAAGGACAAGATA 48
Qy 45 ThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPheArgAspSerSer 64
Db 49 ACAGCACAGGAATATGCAGTAAAGTAATATCTAAACGCTCAAGTAAACACAGAAGACAGAT 108
Qy 65 LeuGlu-----AsnGluIleAlaValLeuLysLysIleLysHisGluAsnIleVal 81
Db 109 AAAGAATTATTTAAAGAGAGTTGAATTATTAAGAATAATTAGATCATCTTAATATCATG 168
Qy 82 ThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuValSer 101
Db 169 AAATTATATGAATCTTTGAGGATAAAGGATACTTTTATCTTTGTACAGAAAGTATATACA 228
Qy 102 GlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAlaSer 121
Db 229 GGAGGAGAATTATTTGATGAAATTAATCGAAAAAGATTTCAGGAGCGCGGATGCAGCT 288
Qy 122 LeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleValHis 141
Db 289 CGTATAGTACGTCAGGTTCTATCGGTATTAATTTATATGTCATCGTAATAAATATGTTTCA 348
Qy 142 ArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIleMet 161
Db 349 AGAGATTTAAAGCCAGAGAATTATTTATTAGAGAATAAAAAAAGATGCAATATACGA 408

```
QY 162 IleThrAspPheGlyLeuSer---LysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 409 ATTATTGATTTCGGGTATCTACACATTTTGAGCCCAAAAAAATGAAGGATAAAATC 468
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 469 GGGACCGGTACTACATGCCCTGAGGTGCTG---CACGGAAACATACGATGAGAAATCC 525
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220
Db 526 GACGTCTGGTCTACGGGTGTATCCCTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 585
QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Db 586 GGAGCAAAATCAATTTGAATTTCTAAGAAAGTCGAGAAAGAAATTCACCTTCGATTTA 645
QY 241 ProPheTrpAspAspLysSerGluSerAlaLysAspPheIleCysHisLeuLeuLys 260
Db 646 CCACAGTGGCGTAAGGTTAGCGAGCCAGCAAAAGATTTAATTAGGAAGATGTTAGCATAT 705
QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTyrIleAspGlyAsn 280
Db 706 GTACCTCAATGCGTATACAGAAAGATGATAGATCATCCATGATGATGATGATGATGAT 765
QY 281 ThrAlaLeuHisArgAspIle-----TyrProSerValSerLeuGlnIleGlnLysAsn 298
Db 766 GATGTTACTGCTAAGATAGTATTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 822
QY 299 PheAlaLysSerTrpArgGlnAlaPheAsnAlaAlaValAlaValHisHisMetArg 318
Db 823 ATCAGGCGATTCAGCGGTACACAGAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882
QY 319 LysLeuHisMetAsn 323
Db 883 AAATTAAACAAAT 897

RESULT 14
US-08-338-057-1
; Sequence 1, Application US/08338057
; Patent No. 5755741
; GENERAL INFORMATION:
; APPLICANT: Tomley, Fiona M.
; APPLICANT: Dunn, Paul P. J.
; APPLICANT: Bumstead, Janene M.
; APPLICANT: Vermeulen, Arno N.
; TITLE OF INVENTION: Coccidiosis poultry vaccine
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Organon Teknika Corporation
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/338,057
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93.309078.9
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 1400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: C-terminal
; ORIGINAL SOURCE:
; ORGANISM: Eimeria maxima
; STRAIN: Houghton
; DEVELOPMENTAL STAGE: sporozoite
; IMMEDIATE SOURCE: sporozoite cDNA cloned in Lambda ZAPII
; LIBRARY: sporozoite cDNA
; CLONE: Em70-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1368
; US-08-338-057-1

Alignment Scores:
Pred. No.: 1.14e-42 Length: 1400
Score: 537.50 Matches: 122
Percent Similarity: 57.70% Conservative: 54
Best Local Similarity: 40.00% Mismatches: 116
Query Match: 21.39% Indels: 13
DB: 1 Gaps: 6

US-09-960-643-2 (1-476) x US-08-338-057-1 (1-1400)
QY 25 PheMetGluValLeuSerGlyAlaPheSerGluValPheLeuValLysGlnArgLeu 44
Db 4 TTCGTGGAAGTTTGGGT-----GAGGTCACTATGCAAGGACAGATA 48
QY 45 ThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPheArgAspSer 64
Db 49 ACAGCACAGGAATATGCAGTAAAGTAATATCTAAACGTCAAGTAAACAGAGACAGAT 108
QY 65 LeuGlu-----AsnGluIleAlaValLeuLysLysIleLysHisGluAsnIleVal 81
Db 109 AAAGATTTATTATTAAAGAAAGTTGAATTTATTAAAGAAATTTAGATCATCTATATCATG 168
QY 82 ThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuValSer 101
Db 169 AATTATATGAATCTTTGAGGATAAAGGATCTTTTATCTGTGTACAGAGATATATACA 228
QY 102 GlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAlaSer 121
Db 229 GGAGGAGAATTTATTTGATGAATTTATTAAATCGAAAAAGATTCAGCGAGCGGATGAGCT 288
QY 122 LeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleValHis 141
Db 289 CGTATAGTACGTACGTTCTATCGGGTATAAATATATATGATCATCTATAAATAATAGTTAT 348
QY 142 ArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGlnAsnSerLysIleMet 161
Db 349 AGAGATTTAAAGCCAGAGAAATTTATTATTAGAGATAAATAAAGAAATGCAATATACGA 408
QY 162 IleThrAspPheGlyLeuSer---LysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 409 ATTATTGATTTTGGGTTTATCTACACATTTTGAGCCCCCAAAAAAATGAAGGATAAAATC 468
QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
Db 469 GGGACCGGTACTACATTCCTGAGGTGCTG---CACGGAAACATACGATGAGAAATGC 525
QY 201 AspCysTrpSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220
Db 526 GACGTCTGGTCTACGGGTGTATCCCTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 585
QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
Db 586 GGAGCAAAATGAATTTGAATTTCTAAGAAAGTCGAGAAAGAAATTCACCTTCGATTTA 645
```




GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:04:21 ; Search time 14 seconds

(without alignments)
1410.195 Million cell updates/sec

Title: US-09-960-643-2

Perfect score: 2513

Sequence: 1 MGRKEEDCSWKKQTINR.....VRKSGSHCRAGGTGVCCLIM 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1155.5	46.0	374	1 KCC1_RAT	Q63450 rattus norv
2	1152.5	45.9	370	1 KCC1_HUMAN	Q14012 homo sapien
3	697.5	27.8	469	1 KCC4_MOUSE	P08414 mus musculu
4	694.5	27.6	295	1 KMLC_DICDI	P25323 dictyosteli
5	692.5	27.6	474	1 KCC4_RAT	P13234 rattus norv
6	684	27.2	473	1 KCC4_HUMAN	Q16566 homo sapien
7	651	25.9	414	1 KCC1_EMEFI	Q00771 emericella
8	643.5	25.6	424	1 KPSH_HUMAN	P11801 homo sapien
9	630	25.1	335	1 KCC1_SCHPO	Q9P712 schizosacch
10	612	24.4	740	1 DCK1_HUMAN	O15075 homo sapien
11	606	24.1	756	1 DCK1_MOUSE	Q9J1M8 mus musculu
12	603.5	24.0	664	1 KCCB_HUMAN	Q13554 homo sapien
13	602.5	24.0	433	1 DCK1_RAT	O08875 rattus norv
14	602	24.0	478	1 KCCA_HUMAN	Q9UCM7 homo sapien
15	599	23.8	478	1 KCCA_RAT	P11275 rattus norv
16	597	23.8	546	1 CHK2_MOUSE	Q9Z265 mus musculu
17	594.5	23.7	499	1 KCCD_HUMAN	Q13557 homo sapien
18	591.5	23.5	533	1 KCCD_RAT	P15791 rattus norv
19	588	23.4	447	1 KCC2_YEAST	P22517 saccharomyc
20	584	23.2	478	1 KCCA_MOUSE	P11798 mus musculu
21	579.5	23.1	542	1 KCCB_MOUSE	P28652 mus musculu
22	578.5	23.0	542	1 KCCB_RAT	P08413 rattus norv
23	578.5	23.0	543	1 CHK2_HUMAN	O96017 homo sapien
24	578	23.0	472	1 KCCG_HUMAN	Q13555 homo sapien
25	577.5	23.0	446	1 KCC1_YEAST	P27466 saccharomyc
26	574	22.8	529	1 KCCG_MOUSE	Q923T9 mus musculu
27	566.5	22.5	527	1 KCCG_RAT	P11730 rattus norv
28	565.5	22.5	382	1 KCC1_METAN	O14408 metarhizium
29	558	22.2	821	1 SPK1_ARATH	P22216 saccharomyc
30	556.5	22.1	610	1 CDPI_YEAST	Q06850 arabidopsis
31	543.5	21.6	508	1 DCPK_SOYBN	P28583 glycine max
32	543.5	21.6	1431	1 DAPK_HUMAN	P53355 homo sapien
33	536	21.3	512	1 RCK1_YEAST	P38522 saccharomyc

RESULT 1

KCC1_RAT

ID	KCC1_RAT	STANDARD;	PRT;	374 AA.
AC	Q63450; Q63084;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Calcium/calmodulin-dependent protein kinase type I (EC 2.7.1.123)			
DE	(CAM kinase I).			
GN	CAMK1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A. (37 KDA ISOFORM).			
RX	MEDLINE=94075341; PubMed=8253780;			
RA	Picciotto M.R., Czernik A.J., Nairn A.C.;			
RT	"Calcium/calmodulin-dependent protein kinase I. cDNA cloning and			
RT	identification of autophosphorylation site.;"			
RL	J. Biol. Chem. 268:26512-26521(1993).			
[2]				
RP	SEQUENCE FROM N.A. (42 KDA ISOFORM).			
RX	STRAIN=Sprague-Dawley; TISSUE=Lung;			
RX	MEDLINE=95035115; PubMed=7948038;			
RA	Cho F.S., Phillips K.S., Boquetti B., Weaver T.E.;			
RT	"Characterization of a rat cDNA clone encoding calcium/calmodulin-			
RT	dependent protein kinase I.;"			
RL	Biochim. Biophys. Acta 1224:156-160(1994).			
[3]				
RN	X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF SHORT FORM.			
RP	TISSUE=Brain;			
RX	MEDLINE=96182648; PubMed=8601311;			
RX	Goldberg J., Nairn A.C., Kuriyan J.;			
RA	"Structural basis for the autoinhibition of			
RT	calcium/calmodulin-dependent protein kinase I.;"			
RL	Cell 84:875-887(1996).			
CC	-1- FUNCTION: PHOSPHORYLATES SYNAPSIN I.			
CC	-1- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein.			
CC	-1- ENZYME REGULATION: ACTIVATED BY CA++/CALMODULIN. MUST BE			
CC	PHOSPHORYLATED TO BE MAXIMALLY ACTIVE.			
CC	-1- SUBUNIT: MONOMER.			
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	CAMK SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L24907; AAA19670.1; -			
DR	EMBL; L26288; AAA66944.1; -			
DR	PDB; 1A06; 08-APR-98.			

P53683 oryza sativ
Q15418 homo sapien
Q28824 bos taurus
Q63531 rattus norv
P29294 oryctolagus
P49101 zea mays (m
P18653 mus musculu
P53684 oryza sativ
P11799 gallus gall
P39009 saccharomyc
Q15746 homo sapien
P53682 oryza sativ

Db 325 -----QEQGQTASHGCELLTPVAGGP-----AAG-C-CRDC----- 354

QY 430 KSSVCSEP 437

Db 355 -----CSEP 358

RESULT 3

KCC4_MOUSE STANDARD; PRT; 469 AA.

AC P08414; Q61381;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Calcium/calmodulin-dependent protein kinase type IV catalytic chain

DE (SC 2.7.1.123) (CAM kinase-GR) (CaMK IV).

GN CAMK4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c; TISSUE=Brain;

RX MEDLINE=91372388; PubMed=1893997;

RA Jones D.A., Glod J., Wilson-Shaw D., Hahn W.E., Sikela J.M.;

RT "cDNA sequence and differential expression of the mouse

RT Ca2+/calmodulin-dependent protein kinase IV gene.";

RL FEBS Lett. 289:105-109(1991).

RN [2]

RP SEQUENCE OF 240-469 FROM N.A.

RX MEDLINE=89122027; PubMed=2536634;

RA Sikela J.M., Law M.L., Kao F.-T., Hartz J.A., Wei Q., Hahn W.E.;

RT "Chromosomal localization of the human gene for brain

RT Ca2+/calmodulin-dependent protein kinase type IV.";

RL Genomics 4:21-27(1989).

RN [3]

RP SEQUENCE OF 315-469 FROM N.A.

RX TISSUE=Brain; PubMed=3033675;

RX MEDLINE=87204263; PubMed=3033675;

RA Sikela J.M., Hahn W.E.;

RT "Screening an expression library with a ligand probe: isolation and

RT sequence of a cDNA corresponding to a brain calmodulin-binding

RT protein.";

PL Proc. Natl. Acad. Sci. U.S.A. 84:3038-3042(1987).

CC -1- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein.

CC -1- SUBCELLULAR LOCATION: SUBSTANTIAL LOCALIZATION IN CERTAIN NEURONAL

CC NUCLEI.

CC -1- TISSUE SPECIFICITY: BRAIN AND TESTIS.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC CAMK SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; M16206; AAA39933.1; -

DR EMBL; M64266; AAA37364.1; -

DR EMBL; J03057; AAA37366.1; -

DR EMBL; X58995; CAA41741.1; -

DR PIR; A29878; A29878.

DR PIR; S17656; S17656.

DR HSSP; Q63450; 1A06.

DR MGD; MGI:88258; Camk4.

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Euk_pkinase; 1.

DR SMART; SM00220; S_TKC; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

KW Transferase; Serine/threonine-protein kinase; ATP-binding;

KW Calmodulin-binding.

FT DOMAIN 42 296 PROTEIN KINASE.

FT NP_BIND 48 56 ATP (BY SIMILARITY).

FT BINDING 71 71 ATP (BY SIMILARITY).

FT ACT_SITE 160 160 BY SIMILARITY.

FT DOMAIN 318 337 CALMODULIN-BINDING (POTENTIAL).

FT CONFLICT 278 280 VLD -> CFGI (IN REF. 2).

FT CONFLICT 302 302 N -> T (IN REF. 2).

SQ SEQUENCE 469 AA; 52627 MW; CE1F98670822F975 CRC64;

Query Match 27.8%; Score 697.5; DB 1; Length 469;

Best Local Similarity 42.7%; Pred. No. 8.6e-37;

Matches 152; Conservative 60; Mismatches 121; Indels 23; Gaps 7;

QY 9 CSSMKKQTTNTRKTF-----IFMEV---LGSQAFSEVFLVKORLTGKLFALKCI 54

Db 14 CSSVASTENLVDPYWDGSDNRDPLGDFEVESELGRGATSVIVRCQKQTKPYALKVL 73

QY 55 KSPAFRDSSLENEIATLVKIKHENIVTLDIYESTTHYLVLMQVSGGELDFRLERGV 114

Db 74 KKT--VDKKIVRTEIGVLLRLSHPNIIKKEIFETPTETISLVLELVGTGGELEFDRIVEKGY 131

QY 115 YTEKDAISLVIOQVLSAVKYLHENGIVHRDLKAPENLLYLTPEENSKIMITDFGLSK-MEQN 173

Db 132 YSERDARDAVRQILEAVAYLHENGIVHRDLKAPENLLYATPAPADAPLKIADFGLSKIYEHQ 191

QY 174 GIMSTACGTPGVVAPEVLAQKPYSKAVDCWSIGVITYILLGYPPEYEET-ESKLFPEKIK 232

Db 192 VLMKTVCTPGYCAPEILRGCAYPEVDMWSVGIIYILLGPEFFDDEGDFQFMRIL 251

QY 233 EGYEYEFSPFWDIDISESAKDFICHLLEKDPNERYTCERKALSHPWIDGNTA--LHRDIYPS 290

Db 252 NCEYVFISFPWDEVSNAKDLVKLLVLDPKKRLTTTQALQHPWVTGKAANFVHMD---T 308

QY 291 VSLQTKNFAKSKWFOAFNAAVHHMKLHMNLHSPGVRPEVENRRPPEQASSETS 346

Db 309 AQKQLQEFNARRKKAAYKAVVASSRLGSSASSSTSTQENHKASSDPPTQDAKDS 364

RESULT 4

KMLC_DICDI

ID KMLC_DICDI STANDARD; PRT; 295 AA.

AC P25323;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin light chain kinase (EC 2.7.1.117) (MLCK).

GN MLKA.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

OX NCBI_TaxID=44689;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AX3;

RX MEDLINE=91340753; PubMed=1651931;

RA Tan J.L., Spudich J.A.;

RT "Characterization and bacterial expression of the Dictyostelium

RT myosin light chain kinase cDNA. Identification of an autoinhibitory

RL domain.";

RL J. Biol. Chem. 266:16044-16049(1991).

RN [2]

RP REVISTONS.

RA Spudich J.A.;

RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.

RN [3]

RP PARTIAL SEQUENCE.

RC STRAIN=AX3;

RX MEDLINE=90337997; PubMed=2380188;

RA Tan J.L., Spudich J.A.;

RT Dictyostelium myosin light chain kinase. Purification and
RT characterization.";
RL J. Biol. Chem. 265:13818-13824(1990).
CC -!- FUNCTION: PHOSPHORYLATES A SPECIFIC SERINE IN THE N-TERMINUS OF A
CC MYOSIN LIGHT CHAIN.
CC
CC -!- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin
CC light-chain] phosphate.
CC -!- ENZYME REGULATION: POSSESSES AN AUTOINHIBITORY DOMAIN.
CC AUTOPHOSPHORYLATION APPEARS TO INCREASE THE ENZYMATIC ACTIVITY.
CC DOES NOT HAVE A CALMODULIN-BINDING DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMK SUBFAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M64176; AAB06337.1; .
DR PIR; A40811; A40811.
DR PIR; A37125; A37125.
DR HSP; Q63450; IA06.
DR DictyDB; DD01034; mlka.
DR InterPro; IPR00719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 8 265 PROTEIN KINASE.
FT NP_BIND 14 22 AUTOINHIBITORY DOMAIN.
FT BINDING 37 37 ATP (BY SIMILARITY).
FT ACT_SITE 130 130 BY SIMILARITY.
SQ SEQUENCE 295 AA; 33406 MW; 546CAEED8F6CDOB CRC64;

Query Match 27.6%; Score 694.5; DB 1; Length 295;
Best Local Similarity 50.2%; Pred. No. 7.6e-37;
Matches 135; Conservative 52; Mismatches 79; Indels 3; Gaps 2;

QY 17 TNIRKTFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAFD--SSLENEI AVLKK 74
DB 2 TEVEKIYEFKEGLRGAFSIVLGENKQTKQRYAIKVINKESELGKDYKELKMEVDILKK 61

QY 75 IKHENIVLTDEYESTHYVLVQLVSGGELFDRLTGRGVYTEKDSLVIOQVLSAVKYL 134
DB 62 VNHPIALKELFDPTEKLYLWELVTGGELFDKIVEKGSSEADAMLVKKIYSAVGYL 121

QY 135 HENGIVHRDLKPNENLYLTPEENSKIMITDFGLSK-MQNGIMSTACGTPGYVAPEVLAQ 193
DB 122 HGLNIVHRDLKPNENLLKSKENHLEVAIADGLSKIIGQTLVMTACGTPSYVAPEVLA 181

QY 194 KPYSAVDCWSGTVITYILLGCPFFYETESKLFPEKTEKGYEFESFPWDISESKADF 253
DB 182 TGYDREVDMSGTVITYILLGCPFFYGDITVPEIPEQIMEANYEPFPEYWGSGISKEARDF 241

QY 254 ICHLEKDPNERYTCCKALSHPWIDGNTA 282
DB 242 ICKLLVDVSKRLNATNALNHPWLKSNNS 270

RESULT 5
KCC4_RAT
ID KCC4_RAT STANDARD; PRT; 474 AA.
AC P13234;
DT 01-JAN-1990 (Rel. 13, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase type IV catalytic chain
DE (EC 2.7.1.123) (CAM kinase-GR) (CAMK IV) (Calspermin).
GN CAMK4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND PARTIAL SEQUENCE.
RX MEDLINE=91288548; PubMed=1648230;
RA Omstede C.-A., Bland M.M., Merrill B.M., Sahyoun N.;
RT "Relationship of genes encoding Ca2+/calmodulin-dependent protein
RT kinase Gr and calspermin: a gene within a gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5784-5788(1991).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=91304387; PubMed=1649385;
RA Means A.R., Cruzalegui F., Lemagueresse B., Needleman D.S.,
RT "A novel Ca2+/calmodulin-dependent protein kinase and a male germ
RT cell-specific calmodulin-binding protein are derived from the same
RT gene.";
RL Mol. Cell. Biol. 11:3960-3971(1991).
RN [3]
RN SEQUENCE OF 250-474 FROM N.A. (ISOFORM 1).
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=89174647; PubMed=2538431;
RA Omstede C.-A., Jensen K.F., Sahyoun N.;
RT "Ca2+/calmodulin-dependent protein kinase enriched in cerebellar
RT granule cells. Identification of a novel neuronal
RT calmodulin-dependent protein kinase.";
RL J. Biol. Chem. 264:5866-5875(1989).
RN [4]
RN SEQUENCE OF 306-474 FROM N.A. (ISOFORM 2), AND SEQUENCE OF 335-361.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=96094352; PubMed=7493991;
RA Sun Z., Means R.L., Lemagueresse B., Means A.R.;
RT "Organization and analysis of the complete rat calmodulin-dependent
RT protein kinase IV gene.";
RL J. Biol. Chem. 270:29507-29514(1995).
CC -!- FUNCTION: CAM KINASE GR IS A NEURONAL-SPECIFIC PROTEIN KINASE,
CC ENRICHED IN CEREELLAR GRANULE CELLS.
CC -!- FUNCTION: CALSPERMIN IS A HEAT-STABLE, ACIDIC, CALMODULIN-BINDING
CC PROTEIN.
CC -!- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein.
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1/calcium-calmodulin-dependent
CC protein kinase type IV catalytic chain (shown here) and
CC 2/calspermin; are produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: Isoform 1 is expressed in brain and isoform 2
CC is testis-specific.
CC -!- PTM: THE N-TERMINAL OF CALSPERMIN IS BLOCKED.
CC -!- MISCELLANEOUS: The presence of an alternative promoter gives rise
CC to the testis-specific isoform 2/calspermin protein.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

```
DR EMBL; M63334; AAA40865.1; -.
DR EMBL; M74488; AAA40845.1; ALT_SEQ.
DR EMBL; M64757; AAA40856.1; -.
DR EMBL; M64757; AAA40857.1; -.
DR EMBL; J04600; AAA41867.1; -.
DR EMBL; J04446; AAA40990.1; -.
DR PIR; A41103; TVRTC4.
DR HSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Calmodulin-binding; Testis; Alternative splicing.
FT DOMAIN 42 296 PROTEIN_KINASE.
FT NP_BIND 48 56 ATP (BY SIMILARITY).
FT BINDING 71 71 ATP (BY SIMILARITY).
FT ACT_SITE 160 160 BY SIMILARITY.
FT DOMAIN 318 337 CALMODULIN-BINDING (POTENTIAL).
FT DOMAIN 393 399 POLY-GLU.
FT DOMAIN 403 413 POLY-GLU.
FT VARSPIC 1 305 MISSING (IN ISOFORM 2).
FT CONFLICT 372 372 I -> M (IN REF. 2 AND 4).
SQ SEQUENCE 474 AA; 53133 MW; 56f71AC5644DED23 CRC64;

Query Match 27.6%; Score 692.5; DB 1; Length 474;
Best Local Similarity 40.8%; Pred. No. 1.8e-36;
Matches 158; Conservative 62; Mismatches 120; Indels 47; Gaps 10;

QY 9 CSWKKQTNIKTF-----IPMEV---LGSGAFSEVFLVKRLTGKLFALKCI 54
Db ||| ||| : : ||| ||| ||| ||| ||| ||| ||| ||| : ||| :
QY 55 KSPAFRDSLSLEAVLKKIKHENIVLEDIESTHYLVMLQVSGSELFDRLLRGV 114
Db || : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 74 KKT--VDKKIVTEIGVLLRSHPNLIKKEIFETPTSLVLEVTGSELFDRIVEGY 131
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 115 YTEKASLVQVLSAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSK-MQN 173
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 132 YSERDAADAVKQILEAVAYLHENGIVHRDLKPENLLYATPADAPLKIADFGSLKIVEHQ 191
QY 174 GIMSTACGTPGVVAPELVLAQKPYKAVDCWSIGVTYILLGYPFFYEET-ESKLFEXIK 232
Db : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : ||| :
QY 192 VLMKTVCGPGPYCAPEILRGCAYGPEVDMWSVGIITYILLGFEFFYDGERGQDFMFRIL 251
QY 233 EGYEFESPFWDIDISEKARDFICHLLEKDPNERYTCERKALSHPWIDGNTA--LHRDIYPS 290
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : ||| :
QY 252 NCEYFIFSPWDEVSINAKDLVKLLVLDPKRLTTFQALQHPWVTGKAANFVHMD---T 308
QY 291 VSLQTKNPAKSKWRQAFNA-----AAVVHMRKRLHMLNLSHSGVVR-----EV 333
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 309 AQKLLQEFNARKRLKAAVAVASSRLGSSASSHTNIOESNKASSEAPQADGDKDTPDL 368
QY 334 ENR-----PPETOASETSRPSPEI 353
QY 369 ENKIQAGDHEAARAAADETMKLOQSEEV 395

RESULT 6
KCC4_HUMAN
ID KCC4_HUMAN STANDARD; PRT; 473 AA.
AC Q16566;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase type IV catalytic chain
DE (EC 2.7.1.123) (CaM kinase-GR) (CaMK IV).
GN CAMK4.
OS Homo sapiens (Human).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94375404; PubMed=8089075;
RA Kitani T., Okuno S., Fujisawa H.;
RT "cDNA cloning and expression of human calmodulin-dependent protein
RL kinase IV.";
RL J. Biochem. 115:637-640(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Cerebellum, and Thymus;
RX MEDLINE=94252566; PubMed=8194751;
RA Bland M.M., Monroe R.S., Ohmstede C.A.;
RT "The cDNA sequence and characterization of the
RT Ca2+/calmodulin-dependent protein kinase-Gr from human brain and
RL thymus.";
RL Gene 142:191-197(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Blood;
RX MEDLINE=94149862; PubMed=8107230;
RA Mosialos G., Hanissian S.H., Jawahar S., Vara L., Kieff E.,
RA Chatila T.A.;
RT "A Ca2+/calmodulin-dependent protein kinase, Cam kinase-Gr, expressed
RT after transformation of primary human B lymphocytes by Epstein-Barr
RT virus (EBV) is induced by the EBV oncogene LMP1.";
RL J. Virol. 68:1697-1705(1994).
RN [4]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: CALSPERMIN IS A HEAT-STABLE, ACIDIC, CALMODULIN-BINDING
CC PROTEIN.
CC -|- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein.
CC -|- SUBCELLULAR LOCATION: SUBSTANTIAL LOCALIZATION IN CERTAIN NEURONAL
CC NUCLEI (BY SIMILARITY).
CC -|- ALTERNATIVE PRODUCTS: CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE
CC TYPE IV CATALYTIC CHAIN AND CALSPERMIN ARE OBTAINED BY ALTERNATIVE
CC SPLICING.
CC -|- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; D30742; BAA06403.1; -.
CC EMBL; L17000; AAA35639.1; -.
CC EMBL; L24959; AAA18251.1; -.
CC EMBL; BC016695; AAH16695.1; -.
CC HSSP; Q63450; 1A06.
CC Genew; HGNC:1464; CAMK4.
CC MIM; 114080; -.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Calmodulin-binding.
FT DOMAIN 46 300 PROTEIN_KINASE.
FT NP_BIND 52 60 ATP (BY SIMILARITY).
FT BINDING 75 75 ATP (BY SIMILARITY).
```

```
FT ACT_SITE 164 341 BY SIMILARITY.
FT DOMAIN 322 341 CALMODULIN-BINDING (POTENTIAL).
SQ SEQUENCE 473 AA; 51925 MW; 8FEE51E5612326DC CRC64;

Query Match 27.2%; Score 684; DB 1; Length 473;
Best Local Similarity 48.1%; Pred. No. 6.1e-36;
Matches 141; Conservative 53; Mismatches 87; Indels 12; Gaps 6;

Qy 25 FMEV---LGSGAFSEVFLVKQRLTGKLFALKCIKKSPAFRDSLSLENEIAVLKKIHENIV 81
Db 45 FFEVESELGRGATSIIVRCRQKQKQKPYALKVLKKT--VDKIVRTIGVLLRLSHPNII 102
Qy 82 TLEDIYSTTHYLLVMQLVSGGELFDRILRGVYTEKDSALVQVLSAVKYVLYHENGIVH 141
Db 103 KLKEIFETPTISLVLVLTGGELEFDRIVEKGYYSERDAADAVKQILEAVAYLHENGIVH 162
Qy 142 RDLKPENLILYTPPEENSKIMTDFGLSK-MEONGIMSTAGCTPGYVAPEVLAQPKYSKAV 200
Db 163 RDLKPENLILYATPADAPLKIADFLGSLKIVHEQVLMKTVCGTGYCAPEILRGCAYPEV 222
Qy 201 DCWSIGVITYLLCGYPPFYEEET-ESKLFEKIKEGYEFESPEFWDIDISESAKDFICHLLE 259
Db 223 DMWSVGIIITYLLCGFPFFYDERGDQFWRRLNCEYFFISPWDEVSLNAKDLVRLKLIIV 282
Qy 260 KDPNERYTCERKALSHPWIDGNTA--LHRDIYPSVLSQIQKNFAKSKWRQAFNA 310
Db 283 LDPKKRLTTFQALQHPWVTGKAANFVHMD---TAQKKLQEFNARRKLKA AVKA 332

RESULT 7
KCC1_EMENI STANDARD; PRT; 414 AA.
AC Q00771;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase (EC 2.7.1.123) (CMPK).
GN CMKA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92225350; PubMed=1563634;
RA Kornstein L.B., Gaiso M.L., Hammell R.L., Bartelt D.C.;
RT "Cloning and sequence determination of a cDNA encoding Aspergillus
nidelans calmodulin-dependent multifunctional protein kinase.";
RL Gene 113:75-82(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=R153;
RA Subbaramaiah K., Greene V., Bartelt D.C.;
RT "Structure of the cmka gene encoding a CamKII homolog in Emericella
(Aspergillus) nidulans.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + protein -> ADP + O-phosphoprotein.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74120; AAB97502.1; -.
DR EMBL; AF054580; AAD22581.1; -.
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
```

```
DR pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW Phosphorylation; ATP-binding.
FT DOMAIN 23 278 PROTEIN KINASE.
FT NP_BIND 29 37 ATP (BY SIMILARITY).
FT BINDING 50 50 ATP (BY SIMILARITY).
FT ACT_SITE 142 142 BY SIMILARITY.
FT DOMAIN 291 301 CALMODULIN-BINDING (BY SIMILARITY).
SQ SEQUENCE 414 AA; 46889 MW; 1E8D58A1C0B2F18C CRC64;

Query Match 25.9%; Score 651; DB 1; Length 414;
Best Local Similarity 38.3%; Pred. No. 6e-34;
Matches 149; Conservative 78; Mismatches 126; Indels 36; Gaps 11;

Qy 20 RKTFFIMEVLVIGSGAFSEVFLVKQR--LTGKLFALKCIKKSPAFRDSLSLENEIAVLKKIKH 77
Db 20 KSLYREGRTLGAAGTYG--IVREADCSSGKVAVKIILKRNVRGNRMVYDELDLQKLNH 76
Qy 78 ENIVTLEDIYESTHYLLVMQLVSGGELFDRILRGVYTEKDSALVQVLSAVKYVLYH 137
Db 77 PHIVHFVDVWFESKDKFYIVTQLATGGELFDRICEYKGFTEKDSQTIROVLDAVNYLHOR 136
Qy 138 GIVHRLDKPENLLYTPPEENSKIMTDFGLSKMEQN--GIMSTAGCTPGYVAPEVLAQKP 195
Db 137 NIVHRLDKPENLLYTPROLDLSQVLADFGIAKMLDNPAEVLTSMAGSFGYAAPEVLMKOG 196
Qy 196 YKAVDCWSIGVITYLLCGYPPFYEEETESKLFKEIKEGYEFESPEFWDIDISESAKDFIC 255
Db 197 HGKAVDWSIGVITYLLCGYSPFRSENLTDLIECSRGRVVFHRYKWDVSKDAKDFIL 256
Qy 256 HLEKDPNERYCEKALSHPWIDGNTALHRDIYPSVLSQIQKNFAKSKWRQAFNAAVVH 315
Db 257 SLLQVDPAPQRTSEALKHPWLKGESASDRDLLP-----EIRAYIARSRLKRGIEIKLAN 312
Qy 316 HMRKLMHNLHSPGVPRVENRPP--ETOASETSRRSPSPITITEAPVL--DHSVALPALT 371
Db 313 RIALKMQ-----EDEEDIPSAVDVQASEASDKSG-----LSPFPALSTENSTHPAST 362
Qy 372 QLPQCHGRRTAPGG---RSLNCLVNGSL 397
Db 363 ----GNGE-----SGGTKKRSLSKIARGAI 383

RESULT 8
KPSH_HUMAN STANDARD; PRT; 424 AA.
ID KPSH_HUMAN
AC P11801; O9NY19;
DT 01-OCT-1989 (Rel. 12, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase H1 (EC 2.7.1.37) (PSK-H1).
GN PSKH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20541713; PubMed=11087665;
RA Brede G., Solheim J., Troen G., Prydz H.;
RT "Characterization of PSKH1, a novel human protein serine kinase with
centrosomal, golgi, and nuclear localization.";
RL Genomics 70:82-92(2000).
RN [2]
RP SEQUENCE OF 199-348 FROM N.A.
RX MEDLINE=87092414; PubMed=2948189;
RA Hanks S.K.;
RT "Homology probing: identification of cDNA clones encoding members of
```


DR pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase;
 KW Calmodulin-binding; Phosphorylation.
 FT DOMAIN 31 291
 FT NP_BIND 37 45
 FT ACT_SITE 154 154
 FT MOD_RES 192 192
 FT MUTAGEN 192 192
 FT CONFLICT 115 115
 FT CONFLICT 294 295
 SQ SEQUENCE 335 AA; 38163 MW; 8761BBCFF6882B02 CRC64;
 Query Match 25.1%; Score 630; DB 1; Length 335;
 Best Local Similarity 43.6%; Pred. No. 9.7e-33;
 Matches 133; Conservative 56; Mismatches 104; Indels 12; Gaps 6;
 QY 28 VLGGAFSEFLVKQLTKLEALKCI-KKSPAFRDSSLENEIAVLKKI--KHENIVTLE 84
 DB 36 VLGGGYATVREAVHETNKNKAAMKMMKKKODEFVKNEIAILKRVSEHPNILLHV 95
 QY 85 DIYESTHYLVMLVGGGELEFDRILERGYTEKDAISLVQOVL SAVRYLHENGIVHRDL 144
 DB 96 DPETVNNLYLTETALGTGELFDRIKAGSGFEYDAALMRTTTSVAVYLHNGIVHRDL 155
 QY 145 KPNLLYLPENSKIMITDFGLSKWEQNG---IMTAGCTGGYVAPEVLAQPKYKAVD 201
 DB 156 KPNLLYRSKOPNSDLLIADFGLSHFYEDSQYMLMTAGCTPEYMAPEFRRTGYGKVPD 215
 QY 202 CWSIGVITVLLCGYPFPEETESKLFKEIKGYEYEFSPFWDIDISEAKDFICHLLEKD 261
 DB 216 MWAIGVITVLLSGYTPFPARPQVEVIEAILANETFPDPCWGSIGSETAKDFIKCLEND 275
 QY 262 PNEYTCERKALSHPWIDGNLTALHRIYPSVLSIQIOKNF-AKSKWQAFNAAVVHHMKRL 320
 DB 276 PSKRLTAADALKHPFLSEKRPATSNLLPNV----RENFNARKTFTRTAYNAVRAFTWKKL 331
 QY 321 HNNLH 325
 DB 332 E-NKH 335
 RESULT 10
 ID DCKL_HUMAN STANDARD; PRT; 740 AA.
 AC O15075;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serine/threonine-protein kinase DCAMKL1 (EC 2.7.1.-) (Doublecortin-like and CAM kinase-like 1).
 DE DCAMKL1 OR KIAA0369.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM AS).
 RC TISSUE=Brain;
 RX MEDLINE=97349984; PubMed=9205841;
 RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RA "prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";
 RT DNA Res. 4:141-150(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS AS AND AL).

RC TISSUE=Fetal brain;
 RX MEDLINE=99156863; PubMed=10036192;
 RA Sossey-Alaoui K., Srivastava A.K.;
 RT "DCAMKL1, a brain-specific transmembrane protein on 13q12.3 that is similar to doublecortin (DCX).";
 RL Genomics 56:121-126(1999).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS AS; AL; BS AND BL).
 RX MEDLINE=98419166; PubMed=9747029;
 RA Omori Y., Suzuki M., Ozaki K., Harada Y., Nakamura Y., Takahashi E.-I., Fujiwara T.;
 RT "Expression and chromosomal localization of KIAA0369, a putative kinase structurally related to Doublecortin.";
 RL J. Hum. Genet. 43:169-177(1998).
 RN [4]
 RP TISSUE SPECIFICITY
 RX MEDLINE=99162404; PubMed=10051403;
 RA Matsumoto N., Pilz D.T., Ledbetter D.H.;
 RT "Genomic structure, chromosomal mapping, and expression pattern of human DCAMKL1 (KIAA0369), a homologue of DCX (XLIS).";
 RL Genomics 56:179-183(1999).
 CC -1- FUNCTION: PROBABLE KINASE THAT MAY BE INVOLVED IN A CALCIUM-SIGNALING PATHWAY CONTROLLING NEURONAL MIGRATION IN THE DEVELOPING BRAIN. MAY ALSO PARTICIPATE IN FUNCTIONS OF THE MATURE NERVOUS SYSTEM.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS: TYPE A SHORT (AS) AND LONG (AL) (SHOWN HERE), TYPE B SHORT (BS) AND LONG (BL); ARE PRODUCED BY ALTERNATIVE SPLICING. TYPE A AND TYPE B ISOFORMS DIFFER RESPECTIVELY BY THE PRESENCE OR ABSENCE OF THE DOUBLECORTIN DOMAIN. AN ALTERNATIVE SPLICING OCCURRING IN 3' OF THE MRNA PRODUCES THE LONG INSTEAD OF THE SHORT ISOFORMS.
 CC -1- TISSUE SPECIFICITY: IN FETAL TISSUES, HIGHLY EXPRESSED IN BRAIN, DETECTABLE IN LUNG AND LIVER, BUT NOT IN KIDNEY. IN ADULT TISSUES, EXPRESSED UBQUITOUSLY IN THE BRAIN, DETECTABLE IN THE HEART, LIVER, SPLEEN, THYMUS, PROSTATE, TESTIS, OVARY, SMALL INTESTINE AND COLON. THE TYPE A ISOFORMS SEEM TO BE EXPRESSED PREDOMINANTLY IN FETAL BRAIN WHEREAS TYPE B ISOFORMS ARE EXPRESSED ABUNDANTLY IN BOTH FETAL AND ADULT BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. CAMK SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 DOUBLECORTIN DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB002367; BAA20824.1; -
 CC HSP; Q63450; IA06.
 CC Genew; HGNC:2700; DCAMKL1.
 CC MIM; 604742; -
 CC InterPro; IPR003533; DCX.
 CC InterPro; IPR000719; Euk_pkinase.
 CC InterPro; IPR002290; Ser_thr_pkinase.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF03607; DCX; 2.
 CC ProDom; PD000001; Euk_pkinase; 1.
 CC SMART; SM00537; DCX; 2.
 CC SMART; SM00220; S_TKC; 1.
 CC PROSITE; PS00309; DC; 2.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
 KW Neurogenesis; Alternative splicing.
 FT DOMAIN 57 143
 FT NP_BIND 186 269
 FT DOMAIN 186 269
 FT DOMAIN 288 340
 FT DOMAIN 390 647
 FT NP_BIND 396 404
 FT ATP (BY SIMILARITY).

DE Calcium/calmodulin-dependent protein kinase type II beta chain (EC
DE 2.7.1.123) (CaM kinase II beta chain) (CaM kinase II beta subunit)
DE (CaM-II beta subunit).
GN CAMK2B OR CAMKB OR CAM2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 0; 1; 4; 6 AND 7).
RC TISSUE=Brain;
RX MEDLINE=20317016; PubMed=10858498;
RA Wang P., Wu Y., Zhou T.H., Sun Y., Pei G.;
RT "Identification of alternative splicing variants of the beta subunit
RT of human Ca(2+)/calmodulin-dependent protein kinase II with different
RT activities";
RL FEBS Lett. 475:107-110(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Skeletal muscle;
RA Leddy J.J., Salih M., Tuana B.S.;
RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 0 AND 4).
RC TISSUE=Brain;
RA Li G.Y., Cooper N.G.F.;
RT "Molecular cloning and sequencing of human calcium/calmodulin
RT dependent protein kinase II beta subunit";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 4).
RC TISSUE=Insulinoma;
RA Rechlitz H., Voigt A., Lankat-Buttgereit B., Goke B., Heimberg H.,
RA Nauck M.A., Schlemann U., Schatz H., Pfeiffer A.F.;
RT "Cloning and quantitative determination of the human
RT Ca2+/calmodulin-dependent protein kinase II (CaMK II) isoforms in
RT human beta cells";
RL Diabetologia 43:465-473(2000).
RN [5]
RP SEQUENCE OF 302-603 FROM N.A. (ISOFORM 1).
RX MEDLINE=97214619; PubMed=9060999;
RA Tames R.M., Krystal G.W.;
RT "Identification of novel human tumor cell-specific CaMK-II variants.";
RL Biochim. Biophys. Acta 1355:281-292(1997).
CC -|- FUNCTION: THIS KINASE MAY PLAY A ROLE IN NEUROTRANSMISSION.
CC -|- CATALYTIC ACTIVITY: ATP + protein = ADP + O-phosphoprotein.
CC -|- ENZYME REGULATION: AUTOPHOSPHORYLATION OF CAM-KINASE II PLAYS AN
CC IMPORTANT ROLE IN THE REGULATION OF THE KINASE ACTIVITY.
CC -|- SUBUNIT: COMPOSED OF FOUR DIFFERENT CHAINS: ALPHA, BETA, GAMMA,
CC AND DELTA.
CC -|- ALTERNATIVE PRODUCTS: 7 ISOFORMS; 0/BETA, 1/BETA1/BETA'E, 2/BETA2,
CC 3 (SHOWN HERE), 4/BETA4/BETA'E, 6/BETA6 AND 7/BETA7; ARE PRODUCED
CC BY ALTERNATIVE SPLICING.
CC -|- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF078803; AAD42035.1; -;
DR EMBL; AF081572; AAD42036.1; -;
DR EMBL; AF083419; AAD42038.1; -;
DR EMBL; AF081924; AAD42037.1; -;
DR EMBL; AF140350; AAD42070.1; -;
DR EMBL; U23460; AAC99802.1; -;
DR EMBL; AF112472; AAD03744.1; -;
DR EMBL; AF112471; AAD03743.1; -;

DR EMBL; AJ252236; CAB65120.1; -;
DR EMBL; AJ252237; CAB65121.1; -;
DR EMBL; AJ252238; CAB65122.1; -;
DR EMBL; U50358; AAB18863.1; -;
DR Genes; HGNC:1461; CAMK2B.
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW Phosphorylation; ATP-binding; Neurone; Alternative splicing.
FT DOMAIN 14 272 PROTEIN KINASE.
FT NP_BIND 20 28 ATP (BY SIMILARITY).
FT BINDING 43 43 ATP (BY SIMILARITY).
FT ACT_SITE 136 136 BY SIMILARITY.
FT DOMAIN 291 301 CALMODULIN-BINDING.
FT VARSPLIC 316 316 V -> A (IN ISOFORM 1, ISOFORM 2 AND
FT ISOFORM 4).
FT VARSPLIC 317 340 MISSING (IN ISOFORM 1, ISOFORM 2 AND
FT ISOFORM 4).
FT VARSPLIC 316 408 MISSING (IN ISOFORM 7).
FT VARSPLIC 354 377 MISSING (IN ISOFORM 6).
FT VARSPLIC 354 392 MISSING (IN ISOFORM 2).
FT VARSPLIC 379 393 MISSING (IN ISOFORM 1).
FT VARSPLIC 410 531 MISSING (IN ISOFORM 0, ISOFORM 1, ISOFORM
2, ISOFORM 4, ISOFORM 6 AND ISOFORM 7).
FT VARSPLIC 557 582 MISSING (IN ISOFORM 6).
FT CONFLICT 68 68 L -> V (IN REF. 3).
FT CONFLICT 533 533 K -> N (IN REF. 5).
SQ SEQUENCE 664 AA; 72726 MW; 9E8C82918D986814 CRC64;
Query Match 24.0%; Score 603.5; DB 1; Length 664;
Best Local Similarity 33.2%; Pred. No. 1e-30;
Matches 156; Conservative 68; Mismatches 179; Indels 71; Gaps 12;
Qy 17 TNIRKTFIFMEVLGSGAFSEVFLVKQLTKGLFALKCI--KKSPAFRDSSELEIAVLKK 74
Db 8 TRFTDEQLYEDIGKGAFAVSRRVCVKLCTGHEYAAKIINTKKLSARDHQKLEAREICRL 67
Qy 75 IKHENIVTLEDIYESTHYILVMOLVSGGELFDRLILRGVYTEKDSLVIQQVLSAVKYL 134
Db 68 LKHSNIVRLHDSISEEGFHYLVFDLVGTGGELFEDIVAREYYSEADASHCQQILEAVLHC 127
Qy 135 HENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMQNGIMST---ACGTPGYVAPEVL 191
Db 128 HQMGVVRDLKPENLLASCKGAAVKLADFGLA-IEVQDQQAQWFCFAGTGYLSPEVL 186
Qy 192 AQKPYSKAVDCWSIGVITYILLGYPFPYFETESKLFKEIKEGYFESFWDIDISAK 251
Db 187 RKEAYGPVDIACGVTILYILLVGYPFPFWDQDKLYQQIKAGAYDFSPSPENDTVTPEAK 246
Qy 252 DFICHLLEKDNERYTCESKALSHPWIDGNLTALHRDIYPSVSLQIQNF-AKSKWRQAF-- 308
Db 247 NLINQMLTINPAKRITAEALHKHPWCQSRSTVASMMHRQETVECLKFNARRKLKGAILT 306
Qy 309 -----NAAAVVHMRKLHMLNLSPGVVRPE---VENRPE 339
Db 307 TMLATRNFSVGRQTAPATMSTAASGTTMGLVEQAQSL-LNKKADGVKQPQTNSTKNSAAA 365
Qy 340 TQASSETSRPSS--PEITITEAPV-----LDHSVALP-----ALTQLPQOQHR 379
Db 366 TSPKGTLPAALEPQTTVIHPVDGKIKESSDANTTIEDAKAPRPVDPDILSSVRRSGA 425
Qy 380 R-----PTAPGGRSLNCLVNGSLHISSSL-----VPMHQGSLAAGCCGCCSSCL 423
Db 426 RSRGAPACPSAPFSPSPAPSPRISDILNSVRRSGTGTPAEAGPLSAGPPCLSPAL 481
RESULT 13
DCK1_RAT

KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW Phosphorylation; ATP-binding; Alternative splicing.
FT DOMAIN 13 271
FT NP_BIND 19 27
FT BINDING 42 42
FT ACT_SITE 135 135
FT MOD_RES 286 286
FT DOMAIN 290 300
FT VARSPLIC 328 328
SQ SEQUENCE 478 AA; 54029 MW; 10800A85CAD724BB CRC64;

Query Match 24.0%; Score 602; DB 1; Length 478;
Best Local Similarity 39.1%; Pred. No. 8.4e-31;
Matches 137; Conservative 58; Mismatches 145; Indels 10; Gaps 5;

QY 17 TNIRKTFIMEVLGSAFSEVFLVKORLTGKLFALKCI--KKSPAFRDSLSLENIATVLKK 74
DB 7 TRFTEEYQLFEELGKGAFAFVVRRCVKVLAGEYAAKIINTKKLSARDHOKLEREARICRL 66

QY 75 IKHENIVTLEDIESTHYVLMQVSGGELFDRILRGVYTEKDASLVIOQVLSAVKYL 134
DB 67 LKHPNIVRLHDSISERGHLYLFDLTGTGELFEDIVAREYISEADASHCIIQIILEAVLHC 126

QY 135 HENGIVHRDLKPENLLYLTPEENSKIMITDFGLS---KMEQNGIMSTAGTGGYVAPEVL 191
DB 127 HQMGVVHRDLKPENLLASLKLGAANKVLADFGLAIEVEGEGQAWFGFA-GTPGYLSPEVL 185

QY 192 AOKPYSAVDCWSIGVITYILLGCGPPPEETESKLFKEKIGEVYEFSPWDIDISESAK 251
DB 186 RNDPYGKPDVWACGVYLLVGVPPFDEQHLRYQQIRAGAYDFPSPEDIVTTPK 245

QY 252 DFTCHLLEKDPNERTCEKALSHPWIDNTALHRDIYPSVLSIQIKNF-AKSKWRQAPNA 310
DB 246 DLINKMLTINPSKRITAAEALHPWISHRSTVASCHMRQETVDCCLKFNARRKLKGAILT 305

QY 311 AAVVHHM---RLHNLNLSGVRPEVENPPTQASETSRSPPITITE 357
DB 306 TMLATRNFGSGGKNKSDGVKSESESTNTTIEDFTVRRKQETIKVTE 355

RESULT 15
KCCA-RAT STANDARD; PRT; 478 AA.
AC P11275;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Calcium/Calmodulin-dependent protein kinase type II alpha chain (EC
DE 2.7.1.123) (CaM-kinase II alpha chain) (CaM kinase II alpha subunit)
DE (CaM-kinase II alpha subunit).
GN CAMK2A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=87289722; PubMed=3475713;
RA Lin C.R., Kapiloff M.S., Durgerian S., Tatamoto K., Russo A.F.,
RA Hanson P., Schulman H., Rosenfeld M.G.;
RT "Molecular cloning of a brain-specific calcium/calmodulin-dependent
RT protein kinase."
RL Proc. Natl. Acad. Sci. U.S.A. 84:5962-5966(1987).
RN [2]
RN SEQUENCE OF 132-327 FROM N.A.
RX MEDLINE=87263392; PubMed=3037704;
RA Hanley R.M., Means A.R., Ono T., Kemp B.E., Burgin K.E., Waxham N.,
RA Kelly P.T.;
RT "Functional analysis of a complementary DNA for the 50-kilodalton
RT subunit of calmodulin kinase II."
RL Science 237:293-297(1987).
RN [3]
RN SEQUENCE OF 1-6 FROM N.A.

RX MEDLINE=90115857; PubMed=2153289;
RA Sunyer T., Sahyoun N.;
RT "Sequence analysis and DNA-protein interactions within the 5',
RT flanking region of the Ca2+/calmodulin-dependent protein kinase II
RT alpha-subunit gene."
RL Proc. Natl. Acad. Sci. U.S.A. 87:278-282(1990).
RN [4]
RN SEQUENCE OF 282-299, AND PHOSPHORYLATION OF THR-286.
RX MEDLINE=88320438; PubMed=2842767;
RA Thiel G., Czernik A.J., Gorelick F., Nairn A.C., Greengard P.;
RT "Ca2+/calmodulin-dependent protein kinase II: identification of
RT threonine-286 as the autophosphorylation site in the alpha subunit
RT associated with the generation of Ca2+-independent activity."
RL Proc. Natl. Acad. Sci. U.S.A. 85:6337-6341(1988).
CC -1- FUNCTION: THIS KINASE MAY PLAY A ROLE IN NEUROTRANSMISSION.
CC -1- CATALYTIC ACTIVITY: ATP + protein -> ADP + O-phosphoprotein.
CC -1- ENZYME REGULATION: Autophosphorylation of Thr-286 allows the
CC kinase to switch from a calmodulin-dependent to a calmodulin-
CC independent state.
CC -1- SUBUNIT: COMPOSED OF FOUR DIFFERENT CHAINS: ALPHA, BETA, GAMMA,
CC AND DELTA.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02942; AAA41870.1; -;
DR EMBL; M16960; AAA41855.1; -;
DR EMBL; M29699; AAA40841.1; -;
DR PIR; A31235; A31235.
DR PIR; A30355; A30355.
DR HSSP; Q63450; IA06.
DR InterPro; IPR00719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKCI; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
FT DOMAIN 13 271
FT NP_BIND 19 27
FT BINDING 42 42
FT ACT_SITE 135 135
FT MOD_RES 286 286
FT DOMAIN 290 300
FT CONFLICT 301 301
SQ SEQUENCE 478 AA; 54114 MW; 306F416CCE9B5F62 CRC64;

Query Match 23.8%; Score 599; DB 1; Length 478;
Best Local Similarity 38.9%; Pred. No. 1.3e-30;
Matches 136; Conservative 59; Mismatches 145; Indels 10; Gaps 5;

QY 17 TNIRKTFIMEVLGSAFSEVFLVKORLTGKLFALKCI--KKSPAFRDSLSLENIATVLKK 74
DB 7 TRFTEEYQLFEELGKGAFAFVVRRCVKVLAGEYAAKIINTKKLSARDHOKLEREARICRL 66

QY 75 IKHENIVTLEDIESTHYVLMQVSGGELFDRILRGVYTEKDASLVIOQVLSAVKYL 134
DB 67 LKHPNIVRLHDSISERGHLYLFDLTGTGELFEDIVAREYISEADASHCIIQIILEAVLHC 126

QY 135 HENGIVHRDLKPENLLYLTPEENSKIMITDFGLS---KMEQNGIMSTAGTGGYVAPEVL 191
DB 127 HQMGVVHRDLKPENLLASLKLGAANKVLADFGLAIEVEGEGQAWFGFA-GTPGYLSPEVL 185

```
QY 192 AQPYSKAVDCWSIGVITYILLGYPFYETEESKLEKIKEGYEFESPFWDDISESAK 251
: || || | : || || | : || || | : || || | : || || | : || || | : || || |
Db 186 RKDPYKGPVDLWACGVILYLLVGYPPFWDEQHRLYQQIKAGAYDFPSPEWDTVTPEAK 245
QY 252 DFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIOKNF-AKSKWRQAFNA 310
| | | | | : || || | : || || | : || || | : || || | : || || | : || || |
Db 246 DLINKMLTINPSKRITAAEALKHPWISHRSTVASCWHROETVDCCLKFNARRKLGAILT 305
QY 311 AAVVHHM---RKLHMLHSPGVREVENRPPETQASSETSRPSSPEITITE 357
: | | | | | : || || | : || || | : || || | : || || | : || || |
Db 306 TMLATRNFGGKSGGNKNDGVKESSESTNTTIEDETKVRQEIIKVTE 355
```

Search completed: March 14, 2003, 14:08:58
Job time : 17 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:07:31 ; Search time 18 Seconds
(without alignments)
2542.221 Million cell updates/sec

Title: US-09-960-643-2

Perfect score: 2513

Sequence: 1 MGRKEEDDCSSWKQTNR.....VKASGSSHCRAQTGVCLIM 476

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1155.5	46.0	374	1 S50193	Ca2+/calmodulin-de
2	1152.5	45.9	370	1 S57347	Ca2+/calmodulin-de
3	1004	40.0	348	2 T37321	Ca2+/calmodulin-de
4	880	35.0	310	2 B88640	protein K07A9.2 [i
5	736	29.3	504	2 I56542	calmodulin-binding
6	703	28.0	421	2 T30814	calmodulin-binding
7	697.5	27.8	469	1 S17656	Ca2+/calmodulin-de
8	692.5	27.6	474	1 TVRRC4	Ca2+/calmodulin-de
9	692.5	27.6	502	2 I52637	Ca2+/calmodulin-de
10	691.5	27.5	301	1 A40811	myosin-light-chain
11	684	27.2	413	1 A50336	Ca2+/calmodulin-de
12	651	25.9	414	2 JN0323	Ca2+/calmodulin-de
13	630	25.1	335	2 T50290	calmodulin kinase
14	604.5	24.1	589	2 S68470	Ca2+/calmodulin-de
15	603	24.0	708	2 T23616	hypothetical prote
16	599.5	23.9	516	1 JU0270	Ca2+/calmodulin-de
17	599	23.8	478	1 A30355	Ca2+/calmodulin-de
18	599	23.8	530	2 D44412	Ca2+/calmodulin-de
19	598.5	23.8	509	2 B44412	calmodulin-depende
20	593	23.6	580	2 T40939	probable Ca-calmod
21	592	23.6	556	2 JC5636	Ca2+/calmodulin-de
22	591.5	23.5	533	1 A34366	Ca2+/calmodulin-de
23	588	23.4	447	2 B40896	Ca2+/calmodulin-de
24	588	23.4	478	1 S04365	Ca2+/calmodulin-de
25	583.5	23.2	547	2 T23614	hypothetical prote
26	582.5	23.2	518	1 S43845	Ca2+/calmodulin-de
27	581.5	23.1	518	1 B46619	Ca2+/calmodulin-de
28	579.5	23.1	542	1 A45025	Ca2+/calmodulin-de
29	578.5	23.0	542	1 A26464	Ca2+/calmodulin-de

30	577.5	23.0	446	2 A40896	Ca2+/calmodulin-de
31	573	22.8	554	2 T05476	calcium-dependent
32	569.5	22.7	504	2 T38226	probable serine-th
33	566.5	22.5	527	1 A31908	Ca2+/calmodulin-de
34	559.5	22.3	490	1 S71776	calcium-dependent
35	558	22.2	821	1 A39616	protein kinase RAD
36	556.5	22.1	610	1 A49082	calcium-dependent
37	553	22.0	490	2 T08873	calcium-dependent
38	546.5	21.7	520	2 F85059	probable calcium d
39	544.5	21.7	583	2 H84810	probable calcium d
40	544	21.6	443	2 T33690	hypothetical prote
41	543.5	21.6	319	2 I38138	protein-serine kin
42	543.5	21.6	508	1 A43713	calcium-dependent
43	543	21.6	495	1 S46284	calcium-dependent
44	541.5	21.5	553	1 T02139	calcium-dependent
45	541.5	21.5	591	2 S54788	calcium-stimulated

ALIGNMENTS

RESULT 1

S50193

Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) I - rat

N:Alternate names: CaMKI

C:Species: Rattus norvegicus (Norway rat)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S50193; A49682; A46038

R:Cho, F.S.; Phillips, K.S.; Bogucki, B.; Weaver, T.E.

Biochim. Biophys. Acta 1224, 156-160, 1994

A:Title: Characterization of a rat cDNA clone encoding calcium/calmodulin-dependent

A:Reference number: S50193; MUID:95035115; PMID:7948038

A:Accession: S50193

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <CHO>

A:Cross-references: EMBL: L26288; NID:9439613; PIDN:AAA66944.1; PID:9439614

R:Picciotto, M.R.; Czernik, A.J.; Nairn, A.C.

J. Biol. Chem. 268, 26512-26521, 1993

A:Title: Calcium/calmodulin-dependent protein kinase I. cDNA cloning and identifi

A:Reference number: A49682; MUID:94075341; PMID:8253780

A:Accession: A49682

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-111; G', 113-117, 'R', 119-308, 'R', 310-322, 'HQP', 327, 'T', 329, 'TDS' <PIC>

A:Cross-references: GB:I24907; NID:9406112; PIDN:AAA19670.1; PID:9406113

R:Mochizuki, H.; Ito, T.; Hidaka, H.

J. Biol. Chem. 268, 9143-9147, 1993

A:Title: Purification and characterization of Ca2+/calmodulin-dependent protein kin

A:Reference number: A46038; MUID:93232082; PMID:8386178

A:Accession: A46038

A:Status: preliminary

A:Molecule type: protein

A:Residues: 12-36, 'T' <MOC>

A:Experimental source: cerebrium

A:Note: sequence extracted from NCBI backbone (NCBIP:129927)

C:Superfamily: Ca2+/calmodulin-dependent protein kinase I; protein kinase homology

C:Keywords: ATP; autophosphorylation; calmodulin binding; phosphoprotein; phosphot

F:18-276/Domain: protein kinase homology <KIN>

F:26-34/Region: protein kinase ATP-binding motif

F:293-299/Region: autoinhibitory

F:302-314/Region: calmodulin binding

F:177/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status pred

Query Match 46.0%; Score 1155.5; DB 1; Length 374;

Best Local Similarity 69.6%; Pred. No. 6.9e-44;

Matches 222; Conservative 43; Mismatches 51; Indels 3; Gaps 3;

QY 12 WKQTNTNRKTFIFMVLGSAFSEVFLVKORLTGKLFALKCI-KKSPAFRDSLENEIA 70

DB 10 W-KQAEIRDYDFRDVLGTGAFSEVILADKRTQKLVAIKIAKKEGSEMEIEA 68

QY 71 VLKKIKHENIVTLEDIYESTTHYVLMQLVSGGELFDRILRGVYTEKDSLVIQQVLSA 130

```
Db 69 VLHKIKHPNIVALDDIYESGGHLYLMQLVSGGELFDRIVEKGFYTERDASRLIFQV LDA 128
QY 131 VKYLHENGIVHRDLKPNLLYITPBENSKIMITDFGLSKMEONG-IMSTACGTPGYVAPE 189
Db 129 VKYLDHGLTVHRDLKPNLLYITSLDSDSKIMISDFGLSKMEDPGSVLSTACGTPGYVAPE 188
QY 190 VLAQPKYSKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFESPFWDIDISES 249
Db 189 VLAQPKYSKAVDCWSIGVITAYILLCGYPFPYDENDAKLFEQILKAEYFEDSPYWDIDISDS 248
QY 250 AKDFICHLLEKDPNERYTCEKALSHPWIDGNLTALHRDIYPSVSLQIQKNFAKSKWRQAFN 309
Db 249 AKDFIRHLEKDPKRFTEQALQHPWIAGDTALDKNIHQSVEQIKKNFAKSKWKQAFN 308
QY 310 AAUVVHMRKLNHLHSPG 328
Db 309 ATAVVRHMRKQLGTSQEG 327

RESULT 2
S57347
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) I - human
N:Alternate names: CamKI
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S57347
R:Haribabu, B.; Hook, S.S.; Selbert, M.A.; Goldstein, E.G.; Tomhave, A.M.
EMBO J. 14, 3679-3686, 1995
A:Title: Human calcium-calmodulin dependent protein kinase I: cDNA cloning, domain struc
.
A:Reference number: S57347; MUID:95369239; PMID:7641687
A:Accession: S57347
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-370 <HAR>
A:Cross-references: EMBL:L41816; NID:G790789; PIDN:AAA99458.1; PID:G790790
C:Genetics:
A:Gene: GDB:CAMK1
A:Cross-references: GDB:642249
C:Superfamily: Ca2+/calmodulin-dependent protein kinase I; protein kinase homology
C:Keywords: ATP; autophosphorylation; calmodulin binding; phosphoprotein; phosphotransfe
F:18-276/Domain: protein kinase homology <KIN>
F:26-34/Region: protein kinase ATP-binding motif
F:293-299/Region: autoinhibitory
F:302-314/Region: calmodulin binding
F:177/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status predicted

Query Match 45.9%; Score 1152.5; DB 1; Length 370;
Best Local Similarity 55.8%; Pred. No. 9.2e-44; Mismatches 62; Indels 81; Gaps 9;
Matches 239; Conservative 46;

QY 12 WKQTNRKTRFTFMVGLGSAFSEVFLVKQLRTGLKLFKCI-KKSPAFRDSLSLENEIA 70
Db 10 W-KQADIRDIYDFRDLVLTGAFSEVLAEDRKTQKLVAIKAEALBKSGSMENEIA 68
QY 71 VLKKIKHENIVTLEDIESTHYLYVMQLVSGGELFDRILRGVYTERDASLVIOQVLSA 130
Db 69 VLHKIKHPNIVALDDIYESGGHLYLMQLVSGGELFDRIVEKGFYTERDASRLIFQV LDA 128
QY 131 VKYLHENGIVHRDLKPNLLYITPBENSKIMITDFGLSKMEONG-IMSTACGTPGYVAPE 189
Db 129 VKYLDHGLTVHRDLKPNLLYITSLDSDSKIMISDFGLSKMEDPGSVLSTACGTPGYVAPE 188
QY 190 VLAQPKYSKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFESPFWDIDISES 249
Db 189 VLAQPKYSKAVDCWSIGVITAYILLCGYPFPYDENDAKLFEQILKAEYFEDSPYWDIDISDS 248
QY 250 AKDFICHLLEKDPNERYTCEKALSHPWIDGNLTALHRDIYPSVSLQIQKNFAKSKWRQAFN 309
Db 249 AKDFIRHLEKDPKRFTEQALQHPWIAGDTALDKNIHQSVEQIKKNFAKSKWKQAFN 308
QY 310 AAUVVHMRKLNHLHSPGVRPEVNRPPETQASSETSRSPPSEITITEAPVLDHSAVLP 369
```

```
Db 309 ATAVVRHMRKQLG-----TS----- 324
QY 370 LTQLPCQHGRPTAPGGRSLNCLVNGSLHISSSLVPHHQGSLAAGPCGCCSSCLNIGSKG 429
Db 325 -----QBGQQTASHGELLTPVAGGP-----AAG-C-CCRDC----- 354
QY 430 KSSYCSEP 437
Db 355 ----CVP 358

RESULT 3
T37321
Ca2+/Calmodulin-dependent protein kinase (EC 2.7.1.123) I - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T37321
R:Eto, K.; Takahashi, N.; Kimura, Y.; Masuho, Y.; Arai, K.; Muramatsu, M.; Tokumitsu,
J. Biol. Chem. 274, 22556-22562, 1999
A:Title: Ca2+/calmodulin-dependent protein kinase cascade in Caenorhabditis elegans.
A:Reference number: Z21686; MUID:99357789; PMID:10428833
A:Accession: T37321
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-348 <ETO>
A:Cross-references: EMBL:AB021864; NID:G5672677; PIDN:BAAB2674.1; PID:G5672678
A:Experimental source: strain Bristol N2; embryonic stage
C:Genetics:
A:Note: cmk-1
C:Superfamily: Ca2+/calmodulin-dependent protein kinase I; protein kinase homology
C:Keywords: ATP; autophosphorylation; calmodulin binding; phosphoprotein; phosphotran

Query Match 40.0%; Score 1004; DB 2; Length 348;
Best Local Similarity 57.3%; Pred. No. 2.5e-37;
Matches 185; Conservative 67; Mismatches 69; Indels 2; Gaps 2;

QY 19 IRKTFIFNEVLGSAFSEVFLVKQLR-TGKLFALKCI-KKSPAFRDSLSLENEIAVLKKIK 76
Db 18 IREKYDFRDLVLTGAFSEVFLAESKSDAGQYAVKCIDKKALKGKESENEIKVLRLKR 77
QY 77 HENIVTLEDIESTHYLYVMQLVSGGELFDRILRGVYTERDASLVIOQVLSAVKYLHE 136
Db 78 HNNIVQLEDITVDEKQFVFLVMELVTGGELFDRIVAKSYTEQDASNLIRQLVLEAVGFMD 137
QY 137 NGIVHRDLKPNLLYLTPEENSKIMITDFGLSKMEONGIMSTACGTPGYVAPEVLAQPKY 196
Db 138 NGVHRDLKPNLLYLNQDEDSKIMISDFGLSKTDSGVMTACGTPGYVAPEVLQPKY 197
QY 197 SKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFESPFWDIDISESAKDFICH 256
Db 198 GKAVDVMSIGVITAYILLCGYPFPYDESANLFAQIIKGEYFEDAPYWDQISGSAKDFITH 257
QY 257 LLEKDPNERYTCEKALSHPWIDGNLTALHRDIYPSVSLQIQKNFAKSKWRQAFNAAVVHH 316
Db 258 LMCCDPEARFTCDALSHPWISGNITAYTHDINGVAVHLKSLAKRNWKAYNAALAIQ 317
QY 317 MRKLHMLNHLSPGVRPEVNRPPE 339
Db 318 LQMLRLSSNSNRLQKQASQOQPE 340

RESULT 4
B88640
protein K07A9.2 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-Aug-2001
C:Accession: B88640
R:anonymous, the C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_
```


RESULT 7

S17656
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) IV - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence.revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S17656; A29878; I49571
R:Jones, D.A.; Glod, J.; Wilson-Shaw, D.; Hahn, W.E.; Sikela, J.M.
FEBS Lett. 289, 105-109, 1991
A:Title: cDNA sequence and differential expression of the mouse Ca(2+)/calmodulin-depend
A:Reference number: S17656; MUID:91372388; PMID:1893997
A:Accession: S17656
A:Molecule type: mRNA
A:Residues: 1-469 <JON>
A:Cross-references: EMBL:X58995; NID:g50366; PIDN:CAA41741.1; PID:g50367
R:Sikela, J.M.; Hahn, W.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3038-3042, 1987
A:Title: Screening an expression library with a ligand probe: Isolation and sequence of
A:Reference number: A29878; MUID:87204263; PMID:3033675
A:Accession: A29878
A:Molecule type: mRNA
A:Residues: 315-469 <SIK>
A:Cross-references: GB:M16206; NID:g200360; PIDN:AAA39933.1; PID:g387512
R:Sikela, J.M.; Law, M.L.; Kao, F.
Genomics 4, 21-27, 1989
A:Title: Chromosomal localization of the human gene for brain Ca2+/calmodulin-dependent
A:Reference number: I49571; MUID:89122027; PMID:2536634
A:Accession: I49571
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 250-277, 'CFG', 281-301, 'T', 303-338, 'X', 340-469 <RES>
A:Cross-references: GB:J03057; NID:g192366; PIDN:AAA37366.1; PID:g192367
A:Experimental source: brain
C:Superfamily: Ca2+/calmodulin-dependent protein kinase; protein kinase homology
C:Keywords: alternative splicing; ATP; calmodulin binding; phosphotransferase; serine/th
F:40-296/Domain: protein kinase homology <KIN>
F:48-56/Region: protein kinase ATP-binding motif
F:306-469/Product: calmodulin binding #status predicted <CSP>
F:318-337/Region: calmodulin binding #status predicted
F:71/Active site: Lys #status predicted

Query Match 27.88; Score 697.5; DB 1; Length 469;
Best Local Similarity 42.78; Pred. No. 7.2e-24;
Matches 152; Conservative 60; Mismatches 121; Indels 23; Gaps 7;

Qy 9 CSSMKKQTNIRKTF-----IFMEV---LGSGAFSEVFLVQRQLTGKLFALKCI 54
Db 14 CSSVTASTENLVDPYWDGNSNRDPLGDDFFVESELGRGATSIYVRCQKGTQRPYALKVL 73
Qy 55 KSPAFRDSLENEIAVLKKTKHENIVTLEDIYESTHYLVMLVSGGELFDRLBERGV 114
Db 74 KKT--VDKKIVRTETGIVLLRSLSHPNIIKLEIFETPTETISLVLELVTTGGELFDRIVEKGY 131
Qy 115 YTEKDSASLVIOOVL SAVKYLHENGIVHRDLKPENLLVLTPEENSKIMITDFGLSK-MEQN 173
Db 132 YSERDADAVKQILEAVAYLHENGIVHRDLKPENLLVLTAPDAPLKIADFGLSKIVEHQ 191
Qy 174 GIMSTACGTCGYAPEVLAQRPYKAVDCWSIGVITVILLCGYPFPYEET-ESKLFKIK 232
Db 192 VLMKTVCGTCGYAPEILRGCAYPEVDMMSVGIIITVILLCGFEFFDDEGDFQMFRRIL 251
Qy 233 EGYVEFPFWDIDISGAQFICHLLEKDPNERYTCBKALSHPWGNTA--LHRDIYPS 290
Db 252 NCEYFISPMWDEVSLNAKDLVKLLVLDPKKRLLTTFQALQHPWTKAANFVHMD---T 308
Qy 291 VSLQIKNFASKRQAFNAVAHVHMRKLMHNLHSPGVPEVNRPPETQASSETS 346
Db 309 AQKKLQEFNARRKLLKAAVKAVASSRLGSSASSSSTSIQENHKASSDPPTQDAKDS 364

RESULT 8

TVRVC4
Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) IV - rat

N:Alternate names: Ca2+/calmodulin-dependent protein kinase Gr
N:Contains: calspermin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1993 #sequence.revision 31-Mar-1993 #text_change 11-Jun-1999
C:Accession: A41103; A41237; A32865; A41250; A32035; A60255; I53706
R:Ohmstede, C.A.; Bland, M.M.; Merrill, B.M.; Sahyoun, N.
Proc. Natl. Acad. Sci. U.S.A. 88, 5784-5788, 1991
A:Title: Relationship of genes encoding Ca(2+)/calmodulin-dependent protein kinase Gr
A:Reference number: A41103; MUID:91288548; PMID:1648230
A:Accession: A41103
A:Molecule type: DNA
A:Residues: 47-141, 'NE', 144-474 <OHL>
A:Cross-references: GB:M74488; NID:g203219; PIDN:AAA40845.1; PID:g203220
A:Note: this sequence has been revised in reference A41237
A:Note: part of this sequence was confirmed by sequencing of cDNA to mRNA
R:Ohmstede, C.A.; Bland, M.M.; Merrill, B.M.; Sahyoun, N.
Proc. Natl. Acad. Sci. U.S.A. 88, 9375, 1991
A:Reference number: A41237
A:Accession: A41237
A:Molecule type: DNA
A:Residues: 142-143 <OH2>
A:Cross-references: GB:M63334
A:Note: this is a revision to the sequence from reference A41103
R:Ohmstede, C.A.; Jensen, K.F.; Sahyoun, N.E.
J. Biol. Chem. 264, 5866-5875, 1989
A:Title: Ca(2+)/calmodulin-dependent protein kinase enriched in cerebellar granule ce
A:Reference number: A32865; MUID:89174647; PMID:2538431
A:Accession: A32865
A:Molecule type: mRNA
A:Residues: 250-474 <OH3>
A:Cross-references: GB:J04600; NID:g206172; PIDN:AAA41867.1; PID:g206173
R:Means, A.R.; Cruzalegui, F.; LeMaquerresse, B.; Needleman, D.S.; Slaughter, G.R.; On
Mol. Cell. Biol. 11, 3960-3971, 1991
A:Title: A novel Ca(2+)/calmodulin-dependent protein kinase and a male germ cell-spec
A:Reference number: A41250; MUID:91304387; PMID:1649385
A:Accession: A41250
A:Molecule type: mRNA
A:Residues: 1-371, 'M', 373-408, 'Q', 410-474 <MEA>
A:Cross-references: GB:M64757
R:Ono, T.; Slaughter, G.R.; Cook, R.G.; Means, A.R.
J. Biol. Chem. 264, 2081-2087, 1989
A:Title: Molecular cloning sequence and distribution of rat calspermin, a high affini
A:Reference number: A32035; MUID:89123272; PMID:2914893
A:Accession: A32035
A:Molecule type: mRNA
A:Residues: 306-371, 'M', 373-474 <ONI>
A:Cross-references: GB:J04446; NID:g203642; PIDN:AAA40990.1; PID:g203643
R:Ono, T.; Means, A.R.
Adv. Exp. Med. Biol. 255, 263-268, 1989
A:Title: Calspermin is a testis specific calmodulin-binding protein closely related t
A:Reference number: A60255; MUID:90144189; PMID:2618865
A:Accession: A60255
A:Molecule type: protein
A:Residues: 335-363 <ON2>
A:Note: the amino end of calspermin was blocked
R:Bland, M.M.
Gene 137, 351-352, 1993
A:Title: Identification of alternate 5' untranslated regions in the gene encoding Ca2
A:Reference number: I53706; MUID:94131312; PMID:8299971
A:Accession: I53706
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-37 <RES>
A:Cross-references: GB:L16999; NID:g310086; PIDN:AAA17443.1; PID:g310087
C:Comment: Ca2+/calmodulin-dependent protein kinase IV is enriched in cerebellar gran
and in sperm cells.
C:Superfamily: Ca2+/calmodulin-dependent protein kinase; protein kinase homology
C:Keywords: alternative splicing; ATP; calmodulin binding; phosphotransferase; serine
F:40-296/Domain: protein kinase homology <KIN>
F:48-57/Region: protein kinase ATP-binding motif
F:306-474/Product: calspermin #status predicted <CSP>
F:318-337/Region: calmodulin binding #status predicted
F:71/Active site: Lys #status predicted

Search completed: March 14, 2003, 14:10:05
Job time : 21 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: March 14, 2003, 14:07:45 ; Search time 181 Seconds
(without alignments)
9488.751 Million cell updates/sec
Title: US-09-960-643-1
Perfect score: 2447
Sequence: 1 tggagtgagctcaagcag.....ttttctctataaaaaaaa 2447
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 501302 seqs, 350932545 residues
Total number of hits satisfying chosen parameters: 1002604
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES										
Result No.	Score	%		Length	DB	ID	Description			
		Query	Match							
1	1734.8	70.9	1738	9	US-09-935-464-4		Sequence 4,	Appli		
2	1339.6	54.7	1383	9	US-09-935-464-2		Sequence 2,	Appli		
3	905.6	37.0	157875	9	US-09-935-464-1		Sequence 1,	Appli		
4	504.6	20.6	1772	9	US-10-024-036B-1		Sequence 1,	Appli		
5	502.2	20.5	1074	9	US-10-024-036B-3		Sequence 3,	Appli		
6	499.4	20.4	501	9	US-09-935-464-7		Sequence 7,	Appli		
7	497	20.3	1578	10	US-09-835-788A-6		Sequence 6,	Appli		
8	479	19.6	480	9	US-09-935-464-46		Sequence 46,	Appli		
9	434.2	17.7	467	9	US-09-935-464-47		Sequence 47,	Appli		
10	432.4	17.7	1372	10	US-09-817-181-1		Sequence 1,	Appli		
11	412.6	16.9	470	9	US-09-935-464-48		Sequence 48,	Appli		
12	324.2	13.2	356	9	US-09-935-464-49		Sequence 49,	Appli		
13	319	13.0	319	9	US-09-935-464-50		Sequence 50,	Appli		
14	291	11.9	476	9	US-09-764-868-189		Sequence 189,	App		
c	266.2	10.9	474	9	US-09-796-692-7777		Sequence 7777,	Ap		
15	244.8	10.0	386	9	US-09-935-464-6		Sequence 6,	Appli		
16	243.6	10.0	1503	10	US-09-797-039-3		Sequence 3,	Appli		
17	243.6	10.0	2297	10	US-09-797-039-1		Sequence 1,	Appli		
18	243.6	10.0	3124	12	US-10-153-921-1		Sequence 1,	Appli		
19	243.6	10.0	2979	12	US-10-153-921-1		Sequence 1,	Appli		

Sequence 201, App
Sequence 1, Appli
Sequence 82, Appli
Sequence 6, Appli
Sequence 3, Appli
Sequence 5, Appli
Sequence 4, Appli
Sequence 1, Appli
Sequence 22, App
Sequence 302, App
Sequence 522, App
Sequence 3371, App
Sequence 502, App
Sequence 3, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 177, App
Sequence 1537, App
Sequence 20, Appli
Sequence 6, Appli
Sequence 8, Appli
Sequence 10, Appli
Sequence 210, App
US-09-764-868-201
US-10-116-332-1
US-10-001-835-82
US-10-024-036B-6
US-09-992-481-3
US-09-992-481-5
US-10-024-036B-4
US-10-096-960-1
US-10-044-090-227
US-09-969-708-302
US-09-954-456-522
US-09-880-107-3371
US-09-925-300-502
US-10-217-357-3
US-09-975-326-3
US-09-934-406-3
US-10-217-357-1
US-09-975-326-1
US-09-934-406-1
US-09-731-872-177
US-09-917-800A-1537
US-09-988-462-20
US-09-940-921B-6
US-09-940-921B-8
US-09-940-921B-10
US-09-925-299-210

ALIGNMENTS

RESULT 1
US-09-935-464-4
; Sequence 4, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1738
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-4

Query Match		70.9%;	Score 1734.8;	DB 9;	Length 1738;		
Best Local Similarity		99.9%;	Pred. No. 0;				
Matches 1736;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps
0:							
QY	53	CTTCAACTCTGGAGGCAATGGGTCGAAGGAAGAGATGACTGCAGTTCTCTGGAGAAAC	112				
Db	1	CTTCAACTCTGGAGGCAATGGGTCGAAGGAAGAGATGACTGCAGTTCTCTGGAGAAAC	60				
QY	113	AGACCACCAACATCCGAAACCTTCAATTTTATGGAAGTCTGGGATCAGGAGCTTTCT	172				
Db	61	AGACCACCAACATCCGAAACCTTCAATTTTATGGAAGTCTGGGATCAGGAGCTTTCT	120				
QY	173	CAGAAGTTTCTCGTGAAGCAAAAGACTGCTGGGAAGCTCTTTGCTCTGAAGTGCATCA	232				
Db	121	CAGAAGTTTCTCGTGAAGCAAAAGACTGCTGGGAAGCTCTTTGCTCTGAAGTGCATCA	180				
QY	233	AGAGTCACCTGCTCCGCGGACAGCAGCCCTGGAGAATGAGATTGCTGTGTGAAAAAGA	292				
Db	181	AGAGTCACCTGCTCCGCGGACAGCAGCCCTGGAGAATGAGATTGCTGTGTGAAAAAGA	240				

QY 293 TCAAGCATGAAACATTGTGACCTGGAGGACATCTATGAGAGCACCAACCACTACTACC 352
D 241 TCAAGCATGAAACATTGTGACCTGGAGGACATCTATGAGAGCACCAACCACTACTACC 300
QY 353 TGGTCATGAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCT 412
D 301 TGGTCATGAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCT 360
QY 413 ACACAGAGAAGATGCCAGTCTGGTGATCCAGAGGCTCTTCTGGCAGTGAATACCTAC 472
D 361 ACACAGAGAAGATGCCAGTCTGGTGATCCAGAGGCTCTTCTGGCAGTGAATACCTAC 420
QY 473 ATGAGAATGGCATCTGCACAGAGACTTAAAGCCCGGAAACCTCTTACCTTACCCCTG 532
D 421 ATGAGAATGGCATCTGCACAGAGACTTAAAGCCCGGAAACCTCTTACCTTACCCCTG 480
QY 533 AAGAGAATCTAAGATCATGATGACTGACTTGTGCTGCTCAAGATGGAACAGAAATGGCA 592
D 481 AAGAGAATCTAAGATCATGATGACTGACTTGTGCTGCTCAAGATGGAACAGAAATGGCA 540
QY 593 TCATGTCCACTGCTGCTGGGACCCAGGCTAGCTGGCTCCAGAGAGTGTGCGCCAGAAAC 652
D 541 TCATGTCCACTGCTGCTGGGACCCAGGCTAGCTGGCTCCAGAGAGTGTGCGCCAGAAAC 600
QY 653 CCTACAGCAAGGCTGTGATGCTGGTCCATCGGCGTCAATCACCTACATATTGCTGTG 712
D 601 CCTACAGCAAGGCTGTGATGCTGGTCCATCGGCGTCAATCACCTACATATTGCTGTG 660
QY 713 GATACCCCGCTTATGAAGAAAGGAGTCTAAGCTTTTCGAGNAGATCAAGGAGGCT 772
D 661 GATACCCCGCTTATGAAGAAAGGAGTCTAAGCTTTTCGAGNAGATCAAGGAGGCT 720
QY 773 ACTATGAGTTTGTGATCTCCATCTCGGATGACATTTCTGAGTCAGCCAAAGGACTTTATT 832
D 721 ACTATGAGTTTGTGATCTCCATCTCGGATGACATTTCTGAGTCAGCCAAAGGACTTTATT 780
QY 833 GCACCTGCTTGAGAAGGATCGGAACGAGCGGTACACCTGTGAGAAGGCTTTGAGTCATC 892
D 781 GCACCTGCTTGAGAAGGATCGGAACGAGCGGTACACCTGTGAGAAGGCTTTGAGTCATC 840
QY 893 CTGGATTACGGAACACAGCGCTCCACCGGACATCTACCCATCAGTCAGCTCCAGA 952
D 841 CTGGATTACGGAACACAGCGCTCCACCGGACATCTACCCATCAGTCAGCTCCAGA 900
QY 953 TCCAGAAGAACTTTGCTAAGAGCAAGTGGAGCAAGCTTCAACGACGACGCTGTGTGC 1012
D 901 TCCAGAAGAACTTTGCTAAGAGCAAGTGGAGCAAGCTTCAACGACGACGCTGTGTGC 960
QY 1013 ACCATGAGGAAGCTACACATGAACCTGCACAGCCCGGGGCTCGGCCAGAGGTGGAGA 1072
D 961 ACCATGAGGAAGCTACACATGAACCTGCACAGCCCGGGGCTCGGCCAGAGGTGGAGA 1020
QY 1073 ACAGGCGGCTGAACCTCAAGCTCAGAACTCTAGACCCAGCTCCCTCAGATCAACA 1132
D 1021 ACAGGCGGCTGAACCTCAAGCTCAGAACTCTAGACCCAGCTCCCTCAGATCAACA 1080
QY 1133 TCACGAGGACCTGTCTGTGACACACAGTGTAGCACTCCCTGCCCTGACCCAAATACCCCT 1192
D 1081 TCACGAGGACCTGTCTGTGACACACAGTGTAGCACTCCCTGCCCTGACCCAAATACCCCT 1140
QY 1193 GCCAGATGGCGCGGCGGCTGCTGCTGGTGGAGGTGCTCAACTGCTGTGCTCAATG 1252
D 1141 GCCAGATGGCGCGGCGGCTGCTGCTGGTGGAGGTGCTCAACTGCTGTGCTCAATG 1200
QY 1253 GCTCCCTCCACATACAGCAGCGCTGGTGCCCATCATCAGGGTCCCTGCGCGCGGGG 1312
D 1201 GCTCCCTCCACATACAGCAGCGCTGGTGCCCATCATCAGGGTCCCTGCGCGCGGGG 1260
QY 1313 CCTGTGGCTGTGCTCCAGCTGCCTGAACATTTGGAGCAAAAGGAAAGTCTCCTACTGCT 1372
D 1261 CCTGTGGCTGTGCTCCAGCTGCCTGAACATTTGGAGCAAAAGGAAAGTCTCCTACTGCT 1320

QY 1373 CTGAGCCACACTCTCTCAAAAAGGCGCAACAAAAACAGAACTTCAAGTCGGAGGTGATGG 1432
D 1321 CTGAGCCACACTCTCTCAAAAAGGCGCAACAAAAACAGAACTTCAAGTCGGAGGTGATGG 1380
QY 1433 TACAGTTAAAGCCAGTGGCAGCTCCCACTGCGGCGAGGAGGAGCTGGAGTCTGTCTCA 1492
D 1381 TACAGTTAAAGCCAGTGGCAGCTCCCACTGCGGCGAGGAGGAGCTGGAGTCTGTCTCA 1440
QY 1493 TTATGTGATTCTCTGGAGCTGTGCTATGTCTACTGCTCAAAATTTTCAGGAGACATATTCAACT 1552
D 1441 TTATGTGATTCTCTGGAGCTGTGCTATGTCTACTGCTCAAAATTTTCAGGAGACATATTCAACT 1500
QY 1553 CCTCTGCTCTTCCAAACCTGTGTCTATCCGGCAGAGGAGGAGGAGCAAGTGGAG 1612
D 1501 CCTCTGCTCTTCCAAACCTGTGTCTATCCGGCAGAGGAGGAGGAGCAAGTGGAG 1560
QY 1613 CAGGCTTTACGAGCAGCTTCTGGCCAGAAGCACCAGCTCTGCCAGCGGGGAGCC 1672
D 1561 CAGGCTTTACGAGCAGCTTCTGGCCAGAAGCACCAGCTCTGCCAGCGGGGAGCC 1620
QY 1673 CCTCATAGGAGGCCAGGAGGAGGCCCAAGGCGTAGAAGCTTTGTGAAGCTGTGAGCA 1732
D 1621 CCTCATAGGAGGCCAGGAGGAGGCCCAAGGCGTAGAAGCTTTGTGAAGCTGTGAGCA 1680
QY 1733 GGAGAAGCGGTGCCACCACTTCCAGGCTCTCCCTGACCTGCTCTATGCCCCAC 1790
D 1681 GGAGAAGCGGTGCCACCACTTCCAGGCTCTCCCTGACCTGCTCTATGCCCCAC 1738

RESULT 2

US-09-935-464-2
; Sequence 2, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCH
; FILE REFERENCE: 3322/1H702 US
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-2

Query Match 54.7%; Score 1339.6; DB 9; Length 1383;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1342; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 70 ATGGGTGCAAGGAAGAAGATGATGCTGAGTTCTCTGGAAGAAACAGACCACACATCCCG 129
D 1 ATGGGTGCAAGGAAGAAGATGATGCTGAGTTCTCTGGAAGAAACAGACCACACATCCCG 60
QY 130 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTTTTCCTGGTG 189
D 61 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCTCAGAAGTTTTCCTGGTG 120
QY 190 AAGCAAAAGACTGACTGGGAAAGCTTTTGTCTCTGAAGTGCATCAAGAGTCACCTGCCCTTC 249
D 121 AAGCAAAAGACTGACTGGGAAAGCTTTTGTCTCTGAAGTGCATCAAGAGTCACCTGCCCTTC 180
QY 250 CGGACAGAGCCTGGAGAATGAGATTGCTGTGTTGAAAAAGATCAAGCATGAACATTT 309
D 181 CGGACAGAGCCTGGAGAATGAGATTGCTGTGTTGAAAAAGATCAAGCATGAACATTT 240
QY 310 GTGACCCCTGGAGGACATCTATGAGACCAACCCACTACTACCTGGTGTGATGAGCTTGT 369

APPLICANT: Bandaru, Rajasekhar
TITLE OF INVENTION: 68730 and 69112, Protein Kinase
FILE REFERENCE: MPI2000-521PIR(M)
CURRENT APPLICATION NUMBER: US/10/024,036B
CURRENT FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: 60/258222
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1074
TYPE: DNA
ORGANISM: Homo sapiens
US-10-024-036B-3

Query Match 20.5%; Score 502.2; DB 9; Length 1074;
Best Local Similarity 71.3%; Pred. No. 3e-133;
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2;

QY 70 ATGGTCGAAAGGAAGAAGATGACTGCAGTTCTCTGGAAGAAACAGACACACATCCGG 129
Db 1 ATGGCCGGGAGAACGGCAGAGCAGCTCCTCTGGAAAAGCAAGCTGAAGACATCAAG 60
QY 130 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAACTTTTCCCTGGTG 189
Db 61 AAGATCTTCAGTTCAAAGAGACCTTCGGAACGGGGCCCTTTCGAACTGGTTTAGCT 120
QY 190 AAGCAAGACTGACTGGGAAGCTTTTGTGCTGGAAGTGATC---AAGAAGTCACTGCC 246
Db 121 GAAGAGAAGCAACTGCGAAGCTCTTTGCTGTGAAGTGATCCCTAAGAAGCGCTGAAG 180
QY 247 TTCGGGACAGCAGCTGGAGATGAGATTGCTGTGTTTGAAGAGATCAAGCATGAAC 306
Db 181 GCGAAGAAAGCAGCATAGAGATGAGATGAGCGCTCTGAGAAAGATTGAAGTGAAT 240
QY 307 ATTGTGACCTGGAGACATCTATGAGAGCACCACCACTACTACCTGTGTCATGCAGCTT 366
Db 241 ATTGTTGCCCTGGAAGACATTTATGAAGCCCAATCACTGTACTTGGTCATGACGTG 300
QY 367 GTTCTGTGGGAGCTCTTTGACCGGATCTCTGGAGCGGGGTCTPACACAGAGAAGAT 426
Db 301 GTCTCCGGTGGAGAGCTGTTTGACCGGATAGTGGAGAAGGGTTTTATACAGAGAAGAT 360
QY 427 GCCAGTCTGTGATCAGCAGGCTTTGTCGGCAGTCAATACCTACATCAGATGGCATC 486
Db 361 GCCAGCACTGTATCCGCAAGTCTTGGACGCGGTGTACTATCTCCACAGAAATGGGCATC 420
QY 487 GTCCACAGAGACTTAAAGCCCGAAAACCTGTCTTACCTTTACCCCTGAAGAGAACTCTAAG 546
Db 421 GTCCACAGAGACTCAAGCCCGAAAATCTTGTACTACAGTCAAGATGAGGAGTCCAAA 480
QY 547 ATCATGATCACTGACTTTTGTCTGTCACAGATGGAACAGAATGG---CATCATGTCCACT 603
Db 481 ATAATGATCACTGACTTTGGATTGTCAAAAATGGAGGGCAAAAGGAGATGTGATGTCCACT 540
QY 604 GCCTGTGGACCCAGGCTACGTGCTCCAGAAAGTCTCGGCCAGAAACCCCTACAGCAAG 663
Db 541 GCCTGTGGAACTCCAGGCTATGTCGCTCTGAACTCTCGCCAGAAACCTTACAGCAAA 600
QY 664 GCTGTGGATGCTGGTCCATCGGCGTCATCACTACATATTGCTGTGGATACCCCGG 723
Db 601 GCGGTTGACTGTGTTCCATCGGAGTGATGCGTACATCTTGTCTGCGGCTACCCCTCT 660
QY 724 TTCTATGAAGAAACGGAGTCTAAGCTTTTTCGAGAAGTCAAGGAGGCTTACTAGAGTTT 783
Db 661 TTTTATGATGAAAATGACTCCAAGCTCTTTTGACGATCCCTCAAGSCGGAATATGAGTTT 720
QY 784 GAGTCTCCATTCTGGGATGACATTTCTGAGTCAGCCAAAGCACTTTATTTGCCACTTGCTT 843
Db 721 GACTCTCCCTACTGGGATGACATCTCCGACTCTGCAAAAAGACTTCATTTCGGAACCTGATG 780
QY 844 GAGAAGGATCCGAACAGCGGTACACCTCTGAGAGGGCCTTGAGTCAATCCCTGGATTGAC 903

Db 781 GAGAAGGACCCGATATAAAGATACACGTGTGAGCAGGCAGCTCGGCACCCCATGGATCGCT 840
QY 904 GAAACACAGGGCCCTCCACCGGACATCTACCATCAGTCAGTCAGCCTCCAGATCCAGAAGAAC 963
Db 841 GGTGACACAGCCCTCAACAAAAACATCCACGAGTCCGTGACGGCCAGATCCGGAAGAAC 900
QY 964 TTGCTAAGACGAAGTGGAGGCAAGCCCTTCAACGAGCAGCAGCTGTGGTGCACCATGAGG 1023
Db 901 TTTGCCAAGACAAATGGAGACAAGCATTTAATGCCACGGCCGCTGTGAGACATATGAGA 960
QY 1024 AAGCTACACAT 1034
Db 961 AACTACACT 971

RESULT 6
US-09-935-464-7
; Sequence 7, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSY
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-7

Query Match 20.4%; Score 499.4; DB 9; Length 501;
Best Local Similarity 99.8%; Pred. No. 1.2e-132;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1525 CTGCAATTTTTCAGGAGACATATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCCGG 1584
Db 1 CTGCAATTTTTCAGGAGACATATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCCGG 60
QY 1585 CAGAGGAGGAAGGCAGAGCAAGTGGAGCGGCTTAGCAGGACAGTTTCTGGCCAGAA 1644
Db 61 CAGAGGAGGAAGGCAGAGCAAGTGGAGCGGCTTAGCAGGACAGTTTCTGGCCAGAA 120
QY 1645 GCACCAAGCTCTGCGCAGGGGCGAGCCCTCATAGAGGGCCAGAGGGAGCCCCAAGG 1704
Db 121 GCACCAAGCTCTGCGCAGGGGCGAGCCCTCATAGAGGGCCAGAGGGAGCCCCAAGG 180
QY 1705 CGTAGAACCTTGTGTAAGCTGTGAGCAGGAGAGAGCGGTGCCACAGCTTCCAGGTCTC 1764
Db 181 CGTAGAACCTTGTGTAAGCTGTGAGCAGGAGAGAGCGGTGCCACAGCTTCCAGGTCTC 240
QY 1765 CTTGACCTGCTGCTTCTATGCCACACCCCTACGTCGCGTGGCTGTGTGCAGTGTACGTA 1824
Db 241 CTTGACCTGCTGCTTCTATGCCACACCCCTACGTCGCGTGGCTGTGTGCAGTGTACGTA 300
QY 1825 GATAGCTCTCGCTGGGTCTGTGCTGTTGTCGTAAGGCTTAATGGGCTGGCCAGGCT 1884
Db 301 GATAGCTCTCGCTGGGTCTGTGCTGTTGTCGTAAGGCTTAATGGGCTGGCCAGGCT 360
QY 1885 GTGTCACTTCTCCAAAGCAAAAGCCATATGGAGCATCTACCCAGACTCCCACTCTGCACAC 1944
Db 361 GTGTCACTTCTCCAAAGCAAAAGCCATATGGAGCATCTACCCAGACTCCCACTCTGCACAC 420
QY 1945 ACTCACTCCCACTCTCAAGCTTCAACCTCTTGSCCAGATTGGGCTCAATTAATGTCGTT 2004

Db 421 ACTCACTCCACCTCTCAAGCCTCCAACCTCTTGCCAGATTGGGCTCATTAATGTCGTT 480

QY 2005 GCCTGCCCATCTGCATGAATG 2025
|||||

Db 481 GCCTGCCCATCTGCATGAATG 501

RESULT 7

US-09-835-788A-6

; Sequence 6, Application US/09835788A

; Patent No. US20020077458A1

; GENERAL INFORMATION:

; APPLICANT: NI et al.

; TITLE OF INVENTION: Death Domain-Containing Receptor Polynucleotides, Polypeptides, a

; FILE REFERENCE: PT018P1

; CURRENT APPLICATION NUMBER: US/09/835,788A

; PRIOR FILING DATE: 2001-04-17

; PRIOR APPLICATION NUMBER: PCT/US00/28666

; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/159,585

; PRIOR FILING DATE: 1999-10-18

; PRIOR APPLICATION NUMBER: 60/167,246

; PRIOR FILING DATE: 1999-11-24

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 1578

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-835-788A-6

Query Match 20.3%; Score 497; DB 10; Length 1578;

Best Local Similarity 71.6%; Pred. No. 1.2e-131;

Matches 682; Conservative 0; Mismatches 265; Indels 6; Gaps 2;

QY 88 GATGACTCGAGTTCCTGGGAAGAACAGACACCACCAACATCCGGAACCTTCATTTTATG 147
|||

Db 15 GAGAGCAGCTCCTCTCGGAAAAGCAAGCTGAAGACATCAAGAAGATCTTCGAGTTCAA 74
|||

QY 148 GAAGTCTGGGATCAGGAGCTTTCTCAGAAATTTTCCTGGTGAAGCAAGACTGACTGGG 207
|||

Db 75 GAGACCTTCGGAACCGGGSCCTTTCCGAAGTGGTTTAGCTGAAGAGAAGGCAACTGGC 134
|||

QY 208 AAGCTCTTGGCTCAAGTGCATC---AAGAAGTCACTGCCCTCCGGGACAGAGCCGTG 264
|||||

Db 135 AAGCTCTTGGCTCAAGTGTATCCCTAAGAAGGCGCTGAAGGGAAGGAAAGCAACATA 194
|||||

QY 265 GAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAACATTTGTGACCTCGGAGGAC 324
|||||

Db 195 GAGATGAGATAGCCGTCTCGAAGAAGATTAAAGCATGAAATATTTGGTCCCTGGAAGAC 254
|||||

QY 325 ATCTATGAGAGCACCACCACCTACTACCTGGTTCATCGACGTTGTTCTGTGGGGAGCTC 384
|||||

Db 255 ATTTATGAAGCCCAATCACTGACTTGGTTCATGTCAGCTGGTGTCCGGTGGAGAGCTG 314
|||||

QY 385 TTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAGGATGCCAGTCTGGTATCCAG 444
|||||

Db 315 TTGACCGGATGATGAGAGGGGTTTTATACAGAGAAGGATGCCAGCACTCTGATCCGC 374
|||||

QY 445 CAGGTCCTTCGCGAGTGAATACCTACATGAGATGSCATCGTCCACAGAGACTTAAG 504
|||||

Db 375 CRAAGTCTGGAGCCGTGTACTATCTCCACAGATGGGCGATCGTCCACAGAGACCTCAAG 434
|||||

QY 505 CCCGAAACCTCTTTACCTTACCCTTGAAGAGAATCTTAAGATCATGATCACTGACITTT 564
|||||

Db 435 CCCGAAATCTCTGTACTACAGTCAAGATGAGGAGTCCAAAATAATGATCAGTGACITTT 494
|||||

QY 565 GGTCTGTCCAAGATGGAACAGAAATGG---CATCATGTCCACTGTCCTGTGGAGCCCGAGC 621
|||||

Db 495 GGATTGTCAAAAATGGAGGGCAAGAGAGATGTGATGTCCACTGCTGCGTGGAACTCCAGGC 554
|||||

QY 622 TAGCTGGCTCCAGAAATGCTGTGCCCCAGAAACCTTACAGCAAGGCTGTGGATTGCTGTC 681

Db 555 TATGTCGCTCCTGAAGTCCTCGCCAGAAACCTTTACAGAAAGCCGTTGACTGCTGGTCC 614
|||||

QY 682 ATCGGCGCTCATACATATGCTCTGTGGATACCCCGCTTCTATGAAGAAACGGAG 741
|||||

Db 615 ATCGGAGTGATGCTACATCTGCTCTGGGTACCCCTCTTTTATGATGAATGAC 674
|||||

QY 742 TCTAAGCTTTTCGAGAAGATCAAGGAGGGCTACTATGAGTTTGAGTCTCACTCTCTGGGAT 801
|||||

Db 675 TCCAAGCTCTTTGAGCAGATCCTCAAGGCGGAATATGAGTTTGACTCTCCCTACTGGGAT 734
|||||

QY 802 GACATTCTTGAGTCAGCCCAAGGACTTTATTTGCCACTTGTGAGAAGGATCCGACGAG 861
|||||

Db 735 GACATCTCCGACTCTGAAAAGACTTCATTCTGGAACTGATGGAGAAGACCCGGAATAAA 794
|||||

QY 862 CGGTACACACTGTGAGAAAGGCTTGAGTCACTCCCTCGGATTCACGGAACACAGCGCCCTCCAC 921
|||||

Db 795 AGATACAGCTGTGAGCAGGACGCTCGGACCCCATGGATCGTGGTACACAGCCCTCAAC 854
|||||

QY 922 CGGACATCTACCCATCAGTCAGCTCCAGATCCAGAAAGACTTTGCTAAGACGAAGTGG 981
|||||

Db 855 AAAAACAATCCACGAGTCCGTGAGCGCCAGATCCGGAACAACTTTGCAAGAGCAATGG 914
|||||

QY 982 AGCAAGCCTTCACGACGAGCTGTGTGCACACATGAGGAGCTACACAT 1034
|||||

Db 915 AGCAAGCATTTAATGCCACGCGCTGTGAGACATATGAGAAACTACACCT 967
|||||

RESULT 8

US-09-935-464-46

; Sequence 46, Application US/09935464

; Publication No. US20030027153A1

; GENERAL INFORMATION:

; APPLICANT: Meyer, Joanne

; APPLICANT: Barrington-Martin, Rory

; APPLICANT: Parker, Alexander

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYC

; FILE REFERENCE: 3322/1H702 US1

; CURRENT APPLICATION NUMBER: US/09/935,464

; PRIOR FILING DATE: 2001-08-23

; PRIOR FILING DATE: 2001-01-09

; NUMBER OF SEQ ID NOS: 90

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 46

; LENGTH: 480

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc_feature

; OTHER INFORMATION: n=a or g or c or t/u

US-09-935-464-46

Query Match 19.6%; Score 479; DB 9; Length 480;

Best Local Similarity 99.8%; Pred. No. 8e-127;

Matches 479; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 41 CATCCTCAGAAGCTTCAACTCTGGAGCAATGGGTGCGAAAAGAAAGATGACTGCAAGTT 100
|||||

Db 1 CATCCTCAGAAGCTTCAACTCTGGAGCAATGGGTGCGAAAAGAAAGATGACTGCAAGTT 60
|||||

QY 101 CCTGGAGAAGAACAGACACCACCAATCCGGAACCTTCATTTTTATGGAAGTCTGGGAT 160
|||||

Db 61 CCTGGAGAAGAACAGACACCACCAATCCGGAACCTTCATTTTTATGGAAGTCTGGGAT 120
|||||

QY 161 CAGGAGCTTTCTCAGAAGTTTTCTCTGTGAGCAAGAACTGACTGGGAAGCTCTTTGCTC 220
|||||

Db 121 CAGGAGCTTTCTCAGAAGTTTTCTCTGTGAGCAAGAACTGACTGGGAAGCTCTTTGCTC 180
|||||

QY 221 TGAAGTGCATCAAGAAGTCACTCCCTTCCTGGGAGACAGCAGCCCTGGAGAATGAGATTGCTG 280
|||||

Db 181 TGAAGTGCATCAAGAAGTCACTCCCTTCCTGGGAGACAGCAGCCTGGAGAATGAGATTGCTG 240
|||||

QY 281 TGTTCARAAAGATCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCA 340
Db 241 TGTTCARAAAGATCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCA 300
QY 341 CCCACTACTAGCTGGTGCATGAGCTTTTCTGCTGGGGAGCTTTTGACCGGATCCTGG 400
Db 301 CCCACTACTAGCTGGTGCATGAGCTTTTCTGCTGGGGAGCTTTTGACCGGATCCTGG 360
QY 401 AGCGGGGTGTCTACACAGAAGATGCCAGTCTGTGTGATCCAGCAGTCTTTGTCGGCAG 460
Db 361 AGCGGGGTGTCTACACAGAAGATGCCAGTCTGTGTGATCCAGCAGTCTTTGTCGGCAG 420
QY 461 TGAATACCTACATGAGATGGCATGCTCCACAGAGACTTTAAAGCCCGAAACCTGCTTT 520
Db 421 TGAATACCTACATGAGATGGCATGCTCCACAGAGACTTTAAAGCCCGAAACCTGCTTT 480

RESULT 9
US-09-935-464-47
; Sequence 47, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; APPLICANT: Parker, Alexander
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; TITLE OF INVENTION: DISORDERS SUCH AS SCHIZOPHRENIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 47
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-47

Query Match 17.7%; Score 434.2; DB 9; Length 467;
Best Local Similarity 96.4%; Pred. No. 5.1e-114;
Matches 450; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 41 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTTCGAAAGGAAGAGATGACTGCAGTT 100
Db 1 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTTCGAAAGGAAGAGATGACTGCAGTT 60
QY 101 CCTGGAAGAAACAGACACCAACATCCGAAAACCTTCATTTTATGGAAGTCTGGGAT 160
Db 61 CCTGGAAGAAACAGACACCAACATCCGAAAACCTTCATTTTATGGAAGTCTGGGAT 120
QY 161 CAGAGCTTCTCAGAAGTTTCTGTGTGAAGCAAGACTGACTGGAGCTTTTGCTC 220
Db 121 CAGAGCTTCTCAGAAAGTTTCTGTGTGAAGCAAGACTGACTGGAGCTTTTGCTC 180
QY 221 TGAAGTGCATCAAGAAGTCACTCCCTTCGGGACAGCAGCCCTGGAGATGAGATTCGCTG 280
Db 181 TGAAGTGCATCAAGAAGTCACTCCCTTCGGGACAGCAGCCCTGAGATGAGATTCGCTG 240
QY 281 TGTTCARAAAGATCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCA 340
Db 241 TGTTCARAAAGATCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCA 300
QY 341 CCCACTACTAGCTGGTGCATGAGCTTTTCTGCTGGGGAGCTTTTGACCGGATCCTGG 399
Db 301 CCCACTACTAGCTGGTGCATGAGCTTTTCTGCTGGGGAGCTTTTGACCGGATCCTGG 360
QY 400 GAGCGGGGTGTCTACACAGAAGATGCCAGTCTGTGTGATCCAGCAGGCTTTGTGTCGGCA 459

Db 361 GAGCGGGGTGTCTACACAGAAGATGCCAGNCTGGGTGATCCACANGTCTTTGTCNGCA 420
QY 460 GTGAAATACCTACATGAGAATGGCATGCTCCACAGAGACTTTAAAGCC 506
Db 421 GTGAAATACCTACATGAGAATGGCATGCTCCACAGAGACTTTAAAGCC 467

RESULT 10
US-09-817-181-1
; Sequence 1, Application US/09817181
; Patent No. US20020142427A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001189
; CURRENT APPLICATION NUMBER: US/09/817,181
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Human
US-09-817-181-1

Query Match 17.7%; Score 432.4; DB 10; Length 1372;
Best Local Similarity 73.9%; Pred. No. 3.2e-113;
Matches 563; Conservative 0; Mismatches 196; Indels 3; Gaps 1;

QY 277 GCTGTGTTGAAAAGATCAAGCATGAAACATTTGTGACCTGGAGGACATCTATGAGAGC 336
Db 204 GATGTTCTGGGCACGATCAAGCACCCCAACATTTGTAGCCTGGATGACATCTATGAGAGT 263
QY 337 ACCACCCACTACTACTGCTGTCATGAGCTTTGTTCTGTGGGGAGCTTTTGACCGGATC 396
Db 264 GGGGGCCACCTCTACCTCATGCTGCTGCTGGGTGGGGAGCTTTTGACCGGATTT 323
QY 397 CTGGAGCGGGGTGCTTACACAGAGAAGATGCCAGTCTGGTGTATCCAGCAGGCTTTGTCG 456
Db 324 GTGGAAGAAAGCTTCTACACGAGCGGGAGCCAGCCGCTCATCTCCAGGTGCTGAT 383
QY 457 GCAGTGAATACCTACATGAGAAATGGCATGCTCCACAGAGACTTTAAAGCCCGAAACCTG 516
Db 384 GCTGTGAAATACCTGACCTGAGCTGGGATTTGTACACCGGGATCTCAAGCCAGAGAACTG 443
QY 517 CTTTACCTTACCCCTGAACAGAACTCTAAGATCATGATCAGTCTGCTGTGCTCCAG 576
Db 444 CTGTACTACAGCTGGATGAGACTCCAAATCATGATCTCCGACTTTGGGCTCTCCAG 503
QY 577 ATGGAACAGAAATGGCA---TCATGTCCACTGCTGTGGGACCCAGGCTAGCTGGCTCCA 633
Db 504 ATGAGGAGCCCGGGCAGTGTGCTCTCCAGCCCTGTGGAATCCGGGATAGCTGGCCCT 563
QY 634 GAAGTGTGGCCCGCAAAACCCCTACAGAGAGCTGTGGAATGCTGTGCTCCTCGGCTCATC 693
Db 564 GAAGTGTGGCCCGCAAAACCCCTACAGAGAGCTGTGGAATGCTGTGCTCCTCGGCTCATC 623
QY 694 ACCTACATATTCCTGCTGTGGATACCCCGCTCTATGAAGAACGAGCTCTAAGCTTTTC 753
Db 624 GCCTACATCTTGTCTGTGGGTTTACCCTCTCCCTTCTATGAGAGATGATGCGCAACTCTTT 683
QY 754 GAGAAGATCAAGAGGGCTACTTATGAGTTTGTGCTCCATCTCGGATGACATTTCTGAG 813
Db 684 GAACAGATTTGAAGCGCGAGTACAGTTTGACTCTCTCTTACTGGGACGACATCTCTGAC 743
QY 814 TCAGCCAGAGACTTTTATTTGCCACTTGTGTGAGAGAGATCCGAGAGCGGTGACACCTGT 873
Db 744 TCTGCCAAGATTTTATCCGGCATTGTATCCGGAGAGAGAGAGAGAGAGAGATTCACCTGT 803
QY 874 GAGAAGCCCTTGAGTCTGATCCCTCGGATTGACGGAACACAGCCCTCCACCGGACATCTAC 933

Db 804 GAGCAGCCCTGACGACCCATGGATTGAGGAGATACAGCTCTAGATAAGAAATATCCAC 863
QY 934 CCATCAGTCAGCCTCCAGATCCAGAGAACCTTTGCTAAGCAAGTGGAGGCAAGCCTTC 993
Db 864 CAGTCGGTGAGTCAGCAGATCAGAGAAGCTTTGCCAAGCAAGTGGAGCAAGCCTTC 923
QY 994 AAGCAGCAGCCTGTGTGCACCAATCATGAGGAAGCTACATG 1035
Db 924 AATGCCACGGCTGTGTGGGCACATGAGGAAACTGCAGCTG 965

RESULT 11
US-09-935-464-48
; Sequence 48, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 48
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-48

Query Match 16.9%; Score 412.6; DB 9; Length 470;
Best Local Similarity 96.2%; Pred. No. 7.7e-108;
Matches 452; Conservative 0; Mismatches 14; Indels 4; Gaps 3;

QY 41 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTGGAAAGAGAGATGACTGCAGTT 100
Db 1 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTGGAAAGAGAGATGACTGCAGTT 60
QY 101 CCTGGAAGAAACAGACACCAACAT--CGGAAACCTTTCATTTTATGGAAGTGTGGGA 159
Db 61 CCTGGAAGAAACAGACACCAACATCCCGAAACCTTTCATTTTATGGAAGTGTGGGA 120
QY 160 TCAGGAGCTTTCAGAA--GTTTTCTGTGTAAGCAAAAGTACTGGGAAGCTCTTTG 217
Db 121 TCAGGAGCTTTCAGAAAGTTCCTGTGTGTAAGCAAAAGTACTGGGAAGCTCTTTG 180
QY 218 CTCTGAGTGCATCAGAGATCAGCTTCCGAGCAGACGCTTGGAGATGAGATTG 277
Db 181 CTCTGAGTGCATCAGAGATCAGCTTCCGAGCAGACGCTTGGAGATGAGATTG 240
QY 278 CTGTGTTGAAAAGATCAAGCATGAAACATTTGTACCTGGAGGACATCTATGAGAGCA 337
Db 241 CTGTGTTGAAAAGATCAAGCATGAAACATTTGTACCTGGAGGACATCTATGAGAGCA 300
QY 338 CCACCACACTACTACCTGGTGCATGAGATGCCAGTCTGGTATCCAGCAGCTTTGTCG 456
Db 301 CCACCACACTACTACCTGGTGCATGAGATGCCAGTCTGGTATCCAGCAGCTTTGTCG 360
QY 397 CTGGAGCGGGTCTTACACAGAGAAGATGCCAGTCTGGTATCCAGCAGCTTTGTCG 420
Db 361 CTGGAGCGGGTCTTACACAGAGAAGATGCCAGTCTGGTATCCAGCAGCTTTGTCG 420
QY 457 GCAGTGAATACCTACATGAGATGCATCGTCCACAGAGACTTAAAGCC 506
Db 421 GCAGTGAATACCTACATGAGATGCATCGTCCACAGAGACTTAAAGCC 470

RESULT 12
US-09-935-464-49
; Sequence 49, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 49
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: n=a or g or c or t/u
US-09-935-464-49

Query Match 13.2%; Score 324.2; DB 9; Length 356;
Best Local Similarity 98.0%; Pred. No. 1.2e-82;
Matches 349; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

QY 41 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTGGAAAGAGAGATGACTGCAGTT 100
Db 1 CATCCTCAGAGCTTCAACTCTGGAGCAATGGTGGAAAGAGAGATGACTGCAGTT 60
QY 101 CCTGGAAGAAACAGACACCAACAT--CGGAAACCTTTCATTTTATGGAAGTGTGGGA 159
Db 61 CCTGGAAGAAACAGACACCAACATCCCGAAACCTTTCATTTTATGGAAGTGTGGGA 120
QY 160 TCAGGAGCTTTCAGAA--GTTTTCTGTGTAAGCAAAAGTACTGGGAAGCTCTTTG 217
Db 121 TCAGGAGCTTTCAGAAAGTTCCTGTGTGTAAGCAAAAGTACTGGGAAGCTCTTTG 180
QY 218 CTCTGAGTGCATCAGAGATCAGCTTCCGAGCAGACGCTTGGAGATGAGATTG 277
Db 181 CTCTGAGTGCATCAGAGATCAGCTTCCGAGCAGACGCTTGGAGATGAGATTG 240
QY 278 CTGTGTTGAAAAGATCAAGCATGAAACATTTGTACCTGGAGGACATCTATGAGAGCA 337
Db 241 CTGTGTTGAAAAGATCAAGCATGAAACATTTGTACCTGGAGGACATCTATGAGAGCA 300
QY 338 CCACCACACTACTACCTGGTGCATGAGATGCCAGTCTGGTATCCAGCAGCTTTTGACCGG 393
Db 301 CCACCACACTACTACCTGGTGCATGAGATGCCAGTCTGGTATCCAGCAGCTTTTGACCGG 356

RESULT 13
US-09-935-464-50
; Sequence 50, Application US/09935464
; Publication No. US20030027153A1
; GENERAL INFORMATION:
; APPLICANT: Meyer, Joanne
; APPLICANT: Barrington-Martin, Rory
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING NEUROPSYCHIA
; FILE REFERENCE: 3322/1H702 US1
; CURRENT APPLICATION NUMBER: US/09/935,464
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: US 09/757,300
; NUMBER OF SEQ ID NOS: 90


```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-935-464-50

Query Match      13.0%; Score 319; DB 9; Length 319;
Best Local Similarity 100.0%; Pred. No. 3.6e-81;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 CATCTCAGAGCTTCAACTCTGGAGCAATGGGTGGAAGCAAGATCACTGCAGTT 100
Db 1 CATCTCAGAGCTTCAACTCTGGAGCAATGGGTGGAAGCAAGATCACTGCAGTT 60
QY 101 CCTGGAAGAAACAGACCAACCAATCCCGGAAACCTTCATTTTATGGAAGTCTGGGAT 160
Db 61 CTGGAGNAACAGACCAACCAATCCCGGAAACCTTCATTTTATGGAAGTCTGGGAT 120
QY 161 CAGGAGCTTTCAGAAAGTTTCTGTTGAAGCAAAAGACTGACTGGGAAGCTTTTGCTC 220
Db 121 CAGGAGCTTTCAGAAAGTTTCTGTTGAAGCAAAAGACTGACTGGGAAGCTTTTGCTC 180
QY 221 TGAAGTCATCAAGAGTCACCTGCTTCCGGACAGCAGCTGGAGAATGAGATTGCTG 280
Db 181 TGAAGTCATCAAGAGTCACCTGCTTCCGGACAGCAGCTGGAGAATGAGATTGCTG 240
QY 281 TGTGAAAAAGATCAAGCATGAAAAACATTGTGACCTCGAGGACATCTATGAGAGCACCA 340
Db 241 TGTGAAAAAGATCAAGCATGAAAAACATTGTGACCTCGAGGACATCTATGAGAGCACCA 300
QY 341 CCCACTACTACCTGGTCAAT 359
Db 301 CCCACTACTACCTGGTCAAT 319

RESULT 14
US-09-764-868-189
; Sequence 189, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 189
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-189

Query Match      11.9%; Score 291; DB 9; Length 476;
Best Local Similarity 97.8%; Pred. No. 4.7e-73;
Matches 316; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 1 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCTCGGATCTCTCAGAACCTTCAACT 60
Db 145 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCTCGGATCTCTCAGAACCTTCAACT 204
QY 61 CTGGAGCAATGGTCAAGGAAGAGATGACTGCAGTTCTCTGGAAGAAACAGACCACC 120
Db 205 CTGGAGCAATGGTCAAGGAAGAGAGATGACTGCAGTTCTCTGGAAGAAACAGACCACC 264
QY 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAGTT 180
; Indels 1; Gaps 1;
```

```
Db 265 AACATCCGGAACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAAGTT 324
QY 181 TTCCTGTTGAAGCAAAAGACTGACTGGAGAGCTCTTTGCTCTGAAGTGCATCAGAAGTCA 240
Db 325 TTCCTGTTGAAGCAAAAGACTGACTGGAGAGCTCTTTGCTCTGAAGTGCATCAGAAGTCA 384
QY 241 CCTGCGCTTCGGGACAGCAGCCCTGGAGATGAGA-TTGCTGTGTTGAAAAAG-ATCAAGC 298
Db 385 CCTGCGCTTCGGGACAGCAGCCCTGGAGATGAGANTTGTGTGTTGAAAAAGTTTCAAGC 444
QY 299 ATGAAAAACATTTGTACCTGGAG 321
Db 445 ATGAAAAACATTTGTGAACCTGGG 467

RESULT 15
US-09-796-692-7777/c
; Sequence 7777, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THE
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7777
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (19)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-7777

Query Match      10.9%; Score 266.2; DB 9; Length 474;
Best Local Similarity 73.8%; Pred. No. 5.8e-66;
Matches 350; Conservative 0; Mismatches 121; Indels 3; Gaps 1;
```

Qy	410	TCTACACAGAGAGGATGCCAGTCTGGTGATCCAGCAGGCTTGTGCGCAGTGAAATACC	469
Db	474	TCACACGGAGCGGAGCGCGCCCTCATCTTCAGGTGCTGGATGCTGTGAAATACC	415
Qy	470	TACATGAGAAATGGCATCGTCCACAGAGACTTAAAGCCCGAAAACCTTGCTTTACCTTACCC	529
Db	414	TGCATGCCCTGGGCATTGTACACCGGATCTCAAGCCAGAGAATCTGCTGTACTACAGCC	355
Qy	530	CTGAAGAGAACTCTAAGATCATGACTTGACTTTGGTCTGTCCAGATGGAAACAAGTG	589
Db	354	TGGATGAAGACTCCAAAATCATGATCCGACTTTGGCCTCTCCAAGATGGAGGACCCGG	295
Qy	590	GCA---TCATGTCCACTGCTGTGGACCCAGGCTACGTGGCTCCAGAAGTCCTGGGCC	646
Db	294	GCAGTGTGCTCTCCACCGCTGTGGAACTCCGGGATACGTGGCCCTGAAGTCCTGGGCC	235
Qy	647	AGAAACCTTACAGCAAGGCTGTGGATTGCTGTCCATCGCGCTCATCACTACATATTGC	706
Db	234	AGAAGCCCTACAGCAAGGCTGTGGATTGCTGTCCATAGGTGTCATCGCCTACATCTTGC	175
Qy	707	TCGTGGATACCCCGTTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAAGATCAAGG	766
Db	174	TCTCGGTTTACCCTCCCTTCTATGACGAGAATGATGCCAAACTCTTTGAACAGATTTTGA	115
Qy	767	AGGCTACTATAGTTTGTAGTCTCCATCTGGGATGACATTTCTGAGTCAGCCAAGGACT	826
Db	114	AGCCGAGTACGAGTTTGACTCTCCTTACTGGGACGACATCTGTGACTGCGCAAGATT	55
Qy	827	TTATTGCCACTTGTGTGAGAGGATCCGAACGAGCGGTACACCTGTGAGAAGG	880
Db	54	TCATCCGGCACTTGATGGAGAGGACCCAGAGAAANNNGATTTCACCNCTGAGCAGG	1

Search completed: March 14, 2003, 17:30:10
Job time : 415 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 14:03:50 ; Search time 95 Seconds
(without alignments)
7899.348 Million cell updates/sec

Title: US-09-960-643-1
Perfect score: 2447
Sequence: 1 tggagtggaagctcaagcag.....ttttctcttaaaaaaaaaa 2447

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	418	17.1	1282	US-08-878-989-12	Sequence 12, Appl
2	418	17.1	1282	US-09-272-796-12	Sequence 12, Appl
3	337.2	13.8	362	US-08-905-223-89	Sequence 89, Appl
4	243.6	10.0	3124	US-09-734-030-1	Sequence 1, Appl
5	197.2	8.1	3471	US-08-715-568A-2	Sequence 2, Appl
6	168.6	6.9	2061	US-09-800-960-1	Sequence 1, Appl
7	163.4	6.7	2132	US-09-159-385-3	Sequence 3, Appl
8	163.4	6.7	2132	US-09-186-277-3	Sequence 3, Appl
9	145.8	6.0	1429	US-09-159-385-4	Sequence 4, Appl
10	145.8	6.0	1429	US-09-186-277-4	Sequence 4, Appl
11	135.6	5.5	1349	US-07-951-715A-20	Sequence 20, Appl
12	135.6	5.5	1349	US-08-459-448A-20	Sequence 20, Appl
13	135.6	5.5	1349	US-08-459-595A-20	Sequence 20, Appl
14	135.6	5.5	1349	US-08-459-504B-20	Sequence 20, Appl
15	135.6	5.5	1349	US-08-459-444-20	Sequence 0, Appl
16	135.6	5.5	1349	US-09-547-422-20	Sequence 1, Appl
17	123	5.0	1400	US-08-464-164-1	Sequence 1, Appl
18	123	5.0	1400	US-08-338-057-1	Sequence 1, Appl
19	123	5.0	1400	US-08-668-416-1	Sequence 1, Appl
20	122.4	5.0	4935	US-08-631-097-3	Sequence 3, Appl
21	122.4	5.0	5886	US-08-810-712-9	Sequence 9, Appl
22	122.2	5.0	2637	US-09-735-934A-1	Sequence 1, Appl
23	120.4	4.9	1735	US-09-746-694-3	Sequence 3, Appl
24	120.4	4.9	1858	US-09-529-093A-1	Sequence 1, Appl
25	119	4.9	1776	US-08-655-352-10	Sequence 10, Appl
26	119	4.9	1776	US-09-258-016-10	Sequence 10, Appl
27	119	4.9	1776	US-09-257-825B-10	Sequence 10, Appl

28 114.8 4.7 1333 4 US-09-142-551A-1 Sequence 1, Appl
29 113.4 4.6 2514 3 US-08-655-352-1 Sequence 1, Appl
30 113.4 4.6 2514 4 US-09-258-016-1 Sequence 1, Appl
31 113.4 4.6 2514 4 US-09-257-825B-1 Sequence 1, Appl
32 108.4 4.4 4162 2 US-08-459-448A-26 Sequence 26, Appl
33 108.4 4.4 4162 3 US-08-459-595A-26 Sequence 26, Appl
34 108.4 4.4 4162 3 US-08-459-504B-26 Sequence 26, Appl
35 108.4 4.4 4162 3 US-08-459-444-26 Sequence 26, Appl
36 108.4 4.4 4162 4 US-09-547-422-26 Sequence 26, Appl
37 108.4 4.4 4165 1 US-07-951-715A-26 Sequence 3, Appl
38 107.2 4.4 2549 4 US-09-467-082-3 Sequence 3, Appl
39 106.8 4.4 2374 4 US-09-347-801-3 Sequence 1, Appl
40 106 4.3 8906 2 US-08-826-267-1 Sequence 1, Appl
41 104.6 4.3 2908 4 US-09-930-181-1 Sequence 1, Appl
42 101.8 4.2 2610 2 US-09-212-771-1 Sequence 1, Appl
43 101.8 4.2 2610 3 US-09-091-058-1 Sequence 1, Appl
44 99.6 4.1 1599 3 US-09-256-465-1 Sequence 1, Appl
45 99.6 4.1 1599 4 US-09-167-322-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-878-989-12
; Sequence 12, Application US/08878989
; Patent No. 5885803
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,989
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT06

CLONE: 827431
US-08-878-989-12

Query Match 17.1%; Score 418; DB 2; Length 1282;
Best Local Similarity 65.5%; Pred. No. 2.8e-109;
Matches 626; Conservative 0; Mismatches 326; Indels 3; Gaps 1;

QY 105 GAAGAAACAGACCCACCAACATCCGGAACACCTTCATTTTATGGAAGTGTGGGATCAGG 164
DB 212 GAAGAAACACACGGAGACATCAGCAGCGTCTACGAGATCCGCGAGAGGTCGGCTCGG 271
QY 165 ACCTTTCTCAGAAGTTTCCCTGGTGAAGCAAGACTGACTGGGAGCTCTTCTCTCGAA 224
DB 272 TGCCTTCTCCGAGGTGGTCTGGCCAGAGCGGGCTCCGACACCTCTGTGGCCCTCAA 331
QY 225 GTGCAT---CAAGAAGTCACTGCTTCCGGGACAGAGCCTCGAGAAATGAGATTGCTGT 281
DB 332 GTGCATCCCAAGAGCCCTCCGGGCAAGGAGGCCCTGTGTGGAACGAGATCGCAGT 391
QY 282 GTTGAAGAAGATCAAGCATGAACATTTGACCCCTGGAGACATCTATGAGACCAAC 341
DB 392 GCTCGTAGGATCAGTACACCCCAACATCGTCGCTCTGGAGGATGCCACGAGAGCCCTTC 451
QY 342 CCACCTACTACCTGGTCACTGAGCTGTCTGCTGGGAGCTCTTTGACCGGATCCTGGA 401
DB 452 CCACCTACTACCTGGCCATGGAAGTGGTGACGGTGGCGAGCTGTTCACCGCATCGGA 511
QY 402 GCGGGGTGTCTACACAGAGAAGGATGCCAGTCTGGTGTATCCAGCAGGTCTTTGCGGCG 461
DB 512 GCGGGGTCTTACACAGAGAAGGATGCCAGTCTGGTGTATCCAGCAGGTCTTTGCGGCG 571
QY 462 GAAATACCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 521
DB 572 CTCTCTACTGACAGCCTGGGAGTGTGACCGGGAGCTCAAGCCCAAGAACTCTCTGTA 631
QY 522 CTTTACCCTGAAGAACTCTAAGTATGATGATGATGATGATGATGATGATGATGATG 581
DB 632 TGCCACGCGCTTTGAGGACTCGAAGATCATGCTCTGACTTTGGACTCTCCAAATCCA 691
QY 582 ACAGATGGCATCATGCTCACTGCTGCTGGGACCCAGGCTACGTGGCTCCAGAAGTCT 641
DB 692 GGCTGGGAACATGCTAGGACACCGCTGTGGGACCCCTGGGATATGTGGCCCAAGAGCT 751
QY 642 GSCCCAGAAACCTTACAGAAAGCTGTGGATTTGCTGGTCCATCGGCTCATCATCATCAT 701
DB 752 GGAGCAGAAACCTTACGGAAGAGCCGTAGATGTGTGGCCCTGGCGCTCATCTCTCAT 811
QY 702 ATTGCTCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAGAT 761
DB 812 CCTGCTGTGGGTACCCCGCTTCTACGAGAGAGGACCCCTGAGCTCTTCAGCCAGAT 871
QY 762 CAAGAGGGCTACTATGAGTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGCA 821
DB 872 CTTGAGGCCAGCTATGATGTTGACTTTCTCTGGGATGACATCTCAGAAATCAGCA 931
QY 822 GGACTTTATTTGCCATTTGCTTTGAGAAGATTCGAAACGAGCGGTACACCTGTGAGAAG 881
DB 932 AGACTTTATTTGGCACCTTCTGGAGCGAGACCTTCAAGAGAGTTTCACTGCGCAACAG 991
QY 882 CTTGAGTCATCTCTGGATTGACGAACACAGCGCCCTCCACCGGACATCTACCCATCAG 941
DB 992 CTTGGGACCTTTGGATTCTTTGGGACACAGGCTTTGGCAGGAGCATCTTAGGGTTGT 1051
QY 942 CAGCCTCCAGATCCAGAGAACTTTGCTTAGAGCAAGTGGAGGCAAGCCTTCAACGAGC 1001
DB 1052 CAGTGAGCAGATCCGGAAGACTTTGCTTGACACACTGGAAGGAGCCTTCAATGCCAC 1111
QY 1002 AGCTGTGTGACACACATGAGAAAGCTACACATGAACCTGCACAGCCCGGGCGTC 1056
DB 1112 CTTGTCTCGGCCACATCCGGAAGCTGGGCAAGATCCCAAGGAGGCGAGGGGCC 1166

RESULT 2

US-09-272-796-12
; Sequence 12, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272.796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT06
; CLONE: 827431
US-09-272-796-12

Query Match 17.1%; Score 418; DB 4; Length 1282;
Best Local Similarity 65.5%; Pred. No. 2.8e-109;
Matches 626; Conservative 0; Mismatches 326; Indels 3; Gaps 1;

QY 105 GAAGAAACAGACCCACCAACATCCGGAACACCTTCATTTTATGGAAGTGTGGGATCAGG 164
DB 212 GAAGAAACACACGGAGACATCAGCAGCGTCTACGAGATCCCGAGAGGCTCGGCTCGG 271
QY 165 ACCTTTCTCAGAAGTTTCCCTGGTGAAGCAAGACTGACTGGGAGCTCTTTGCTCTGAA 224
DB 272 TGCCTTCTCCGAGGTGGTCTGGCCAGGAGCGGGCTCCGACACCTCTGTGGCCCTCAA 331
QY 225 GTGCAT---CAAGAAGTCACTGCTTCCGGGACAGAGCCTCGGAAATGAGATTGCTGT 281
DB 332 GTGCATCCCAAGAGCCCTCCGGGCAAGGAGGCCCTGTGTGGAACGAGATCGCAGT 391
QY 282 GTTGAAGAAGATCAAGCATGAACATTTGACCCCTGGAGACATCTATGAGACCAAC 341
DB 392 GCTCGTAGGATCAGTACCCCAACATCGCTCTGGAGGATGTCCACGAGAGAGCCCTTC 451
QY 342 CCACCTACTACCTGGTCACTGAGCTGTGTTCTGGTGGGAGCTCTTTGACCCGATCCTTGA 401

Db 452 CCACCTCTACCTGGCCATGGAACGTGCTGACGGGTGCGGAGCTGTTTGACCGCATCATGGA 511
QY 402 GCGGGGTGCTACACAGAGAGATGCCAGTCTGGTGATCCACGAGGCTTTGTCGGCAGT 461
Db 512 GCGGGGTCTACACAGAGAGATGCCAGGATCTGTTGGTCTGTTGGCGCCGT 571
QY 462 GAAATACCTACATGAGAAATGGCATCTCCACAGAGACTTAAAGCCCGAAACCTGCTTTA 521
Db 572 CTCCTACCTGCACAGCTGGGATCTGCACCGGACCTCAGCCCGAAACCTCTGTA 631
QY 522 CTTTACCCCTGAAGAACTCTAAGATCATGATCATGCTGTTGTTCTTCCAAGATGA 581
Db 632 TGGCAGCCCTTTGAGGACTCGAAGATCATGTTCTCTGACTTTGGACTCTCCAAAATCCA 691
QY 582 ACAGATGGCATCATCTGCACCTGCTGGACCCGACGCTAGTGGCTCCAGAAAGTCT 641
Db 692 GCGTGGAAACATCTAGGACCCCTGTTGGACCCCTGGATATGTTGGCCCGAGAGCTTT 751
QY 642 GGCCCAAGAACCTACAGCAAGCTGTGGATTGCTGGTCCATCGGCGTCACTACCTACAT 701
Db 752 GGAGCAGAAACCTACGGGAAGCCGTAGATGTGTGGGCCCTGGGGTCTATCTCTACAT 811
QY 702 ATTGCTCTGTGATACCCCGTCTTATGAAGAAAGGAGTCTAAGCTTTTCGAGAAGAT 761
Db 812 CTGCTGTGTGGTACCCCGCTTCTACGACGAGAGCGACCTGAGCTCTTCAGCCAGAT 871
QY 762 CAAGGAGGCTACTATGAGTTGAGTCTCCATCTCTGGGATGACATTTCTGAGTCAAGCAA 821
Db 872 CTGAGGGCCAGTATGAGTTGAGTCTTCTTCTGGGATGACATCTCAGAAATCAGGCAA 931
QY 822 GGACTTTATTGCACTTCTTGAGAAAGATCCGAAGAGCGGTACACCTGTGAGAAGGC 881
Db 932 AGACTTTATTGCGACCTTCTGGAGCGAGACCTTCAGAAAGAGTTTCACTGCCAACAGCG 991
QY 882 CTTGAGTCACTCTGAGTGAACGAAACACGGCCCTCCACCGGGGACATCTACCCATCAT 941
Db 992 CTTGCGGGACCTTTGATCTTTTGGGACACAGGCTTTGSCAGGACATCTTAGGGTTGT 1051
QY 942 CAGCCTCCAGATCCAGAAACTTTCTAAGACGATGGAGGAGCGGAGCCCTTCAACGCGAGC 1001
Db 1052 CAGTGAAGATCCGGAAGAACTTTGCTTGGACACACTTGAAGCGAGGCTTCAATGCCAC 1111
QY 1002 AGCTGTGTGCACCATCAGGAAGCTACACATGAACCTGACAGCCCGGGGCTC 1056
Db 1112 CTTGTTCTGGCCACATCCGNAAGCTGGGGCAGATCCAGAGGGGGGGCC 1166

RESULT 3

US-08-905-223-89

; Sequence 89, Application US/08905223

; Patent No. 622029

; GENERAL INFORMATION:

; APPLICANT: Edwards, Jean-Baptiste D.

; APPLICANT: Duclert, Aymeric

; APPLICANT: Lacroix, Bruno

; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS

; NUMBER OF SEQUENCES: 503

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 501 West Broadway

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92101-3505

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: Win95

; SOFTWARE: Word

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/905,223

; FILING DATE:

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned A.

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER:

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 235-8550

; TELEFAX: (619) 235-0176

; INFORMATION FOR SEQ ID NO: 89:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 362 base pairs

; TYPE: NUCLEIC ACID

; STRANDEDNESS: DOUBLE

; TOPOLOGY: LINEAR

; MOLECULE TYPE: CDNA

; ORIGINAL SOURCE:

; ORGANISM: Homo Sapiens

; TISSUE TYPE: Brain

; FEATURE:

; NAME/KEY: sig_peptide

; LOCATION: 87..191

; IDENTIFICATION METHOD: Von Heijne matrix

; OTHER INFORMATION: score 5.6

; OTHER INFORMATION: seq FIFMEVLGSAFs/EV

US-08-905-223-89

Query Match

Best Local Similarity 13.8%; Score 337.2; DB 4; Length 362;

Mismatches 1; Conservative 99.7%; Pred. No. 1.5e-86;

Matches 337; Indels 0; Gaps 0;

QY 1 TGGAGTGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 60

Db 18 TGGAGTGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCCTCAGAGCTTCAACT 77

QY 61 CTGGAGGCAATGGGTGGAAGAGAGATGACTGCAGTTCCTTGGAAAGAACACACACC 120

Db 78 CTGGAGGCAATGGGTGGAAGAGAGATGACTGCAGTTCCTTGGAAAGAACACACACC 137

QY 121 AACATCCGGAACCTTCATTTTATGGAAGTGTGGATCAGGAGCTTCTCAGAAGTT 180

Db 138 AACATCCGGAACCTTCATTTTATGGAAGTGTGGATCAGGAGCTTCTCAGAAGTT 197

QY 181 TTCTGTTGGAAGAAAGACTGACTGGGAAGCTTTTCTCTGTAAGTGCATCAAGACTCA 240

Db 198 TTCTGTTGGAAGAAAGACTGACTGGGAAGCTTTTCTCTGTAAGTGCATCAAGACTCA 257

QY 241 CCTGCCTCCGGGACAGCAGCTGGAGAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 300

Db 258 CCTGCCTCCGGGACAGCAGCTGGAGAATGAGATTGCTGTGTTGAAAAAGATCAAGCAT 317

QY 301 GAAACATTTGACCTGGAGGACATCTATGAGAGCAC 338

Db 318 GAAACATTTGACCTGGAGGACATCTATGAGAGCAC 355

RESULT 4

US-09-734-030-1

; Sequence 1, Application US/09734030

; Patent No. 6461846

; GENERAL INFORMATION:

; APPLICANT: BEASLEY, Ellen M.

; APPLICANT: MERKLOV, Gennady

; APPLICANT: KETCHUM, Karen A.

; APPLICANT: WEI, Ming-Hui

; APPLICANT: DIFRANCESCO, Valentina

; APPLICANT: YAN, Chunhua

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: CLO00612

; CURRENT APPLICATION NUMBER: US/09/734,030

; CURRENT FILING DATE: 2000-12-12

; PRIOR APPLICATION NUMBER: 60/207,281

Db 33 AACGTGGTGGCCCTCCGCGGCGGTACGAGCACAAGCAGACGCTGTCACCTCGTCGGAG 92
Qy 364 CTTGTTCTGTTGGGAGCTCTTTGACCGGATCCTGAGGCGGGTGTCTACACAGAGAAG 423
Db 93 CTGTGCGGCGGCGGGAGCTCTTGACCGCATATCCTCCGCGGCCAGTACACGGAGCGC 152
Qy 424 GATGCACTGTTGTTACCAAGCAGGTCTTGTGCGSCAGTGAATACCTACATGAGATGCC 483
Db 153 GCGCGCGGAGCTGCTGCGGCGCATGTCAGATGTCGACACCTGCCACTCCATGGGG 212
Qy 484 ATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCCCTGAAGAGAATCT 543
Db 213 GTGATGACCGGGACATCAAGCCCGAGAACTTCTGCTGCTCAGCAAGGAGGAGCGG 272
Qy 544 AAGATCATGATCATGCTTGTGTTGCTGCTCAAGATGGAACAGATGGCATGTCCTACT 603
Db 273 CCGCTCAAGGCGCACCGACTTGGGCTCTCCGCTCTTCTTCAAGGAGGCGGAGCTGCTCAGG 332
Qy 604 GCCTGTGGGACCCAGGCTAGTGGCTCCAGAAAGTGTGCGCCAGAAACCCCTACAGCAAG 663
Db 333 GACATGCTGCGCAGCGCTACTACATCGCGCCGAGGTGCTCAAGAGGAAGTACGCCCG 392
Qy 664 GCTGTGATGCTGGTCCATCGGCGTCATCACCTACATATGCTGTGTGGATACCCCGCG 723
Db 393 GAGGCCGACATCTGGAGCGTGGCGTCTGCTGCTACATCTTCTCGCGCGGCTGCTGCC 452
Qy 724 TTCTATGAAGAAAGGAGTCTAAGCTTTTCGAGAAGTCAAGGAGGCTACTATGAGTTT 783
Db 453 TTCTGGCAGAGAACAGACGCGATCTTCACCGGCCATCTCGGAGGCGGAGCTTGACCTC 512
Qy 784 GAGTCTCCATTTCTGGATGACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTGCTT 843
Db 513 TCCAGCGGCGCATGCGCCACACATCTCGCGGGAGCCAGGATCTCGTCARAAGATGCTC 572
Qy 844 GAGAAGGATCGAACGAGCGGTACACCTGTGAGAAGGCTTGTAGTCACTCCCTGGATTGAC 903
Db 573 AACATCAACCCAGAGGCGGCTCACGCGCTTCCAGGCTCTCAATCACCCATGATCAAA 632
Qy 904 GGAAC 909
Db 633 GAAGAC 638

RESULT 13

US-08-459-595A-20
; Sequence 20, Application US/08459595A

; Patent No. 6018104

; GENERAL INFORMATION:

; APPLICANT: Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.

; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED

; NUMBER OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE

; NUMBER OF SEQUENCES: 94

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6018104artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown

; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,595A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8582
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1349 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1226
; OTHER INFORMATION:
; OTHER INFORMATION: pollen-specific calcium dependent protein kinase gene as
; OTHER INFORMATION: disclosed in Figure 30."

US-08-459-595A-20

Query Match

Best Local Similarity 5.5%; Score 135.6; DB 3: Length 1349;

Matches 312; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 304 AACATGTGACCTGGAGGACATCTATGAGGACACCCACCTACTACTGCTGTCATGCG 363
Db 33 AACGTGGTGGGCTCCGCGGCGGTACGAGGACAAGCAGAGCGTGCACCTCGTCATGGAG 92
QY 364 CTTGTTCTGTTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAAG 423
Db 93 CTGTGCGGCGGCGGGAGCTCTTCGACCGCATCTCGCGCGGCCAGTACACGGAGCGC 152
QY 424 GATGCACTGCTGTTGATCCAGCAGGTCTTGTGCGCAGTGAATACCTACATGAGAATGCC 483
Db 153 GCGCGCGGAGCTGCTGCGGCCATCTGTCAGATCGTGCACACCTGCCACTCCATGGGG 212
QY 484 ATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAATCT 543
Db 213 GTGATGACCGGGACATCAAGCCCGAGAACTTCTGCTGCTCAGCAAGGAGGAGCGG 272
QY 544 AAGATCATGATCATGCTTGTGTTGCTGCTCAAGATGGAACAGATGGCATGTCCTACT 603
Db 273 CCGCTCAAGGCGCACCGACTTGGGCTCTCCGCTCTTCTTCAAGGAGGCGGAGCTGCTCAGG 332
QY 604 GCCTGTGGGACCCAGGCTAGTGGCTCCAGAAAGTGTGCGCCAGAAACCCCTACAGCAAG 663
Db 333 GACATGCTGCGCAGCGCTACTACATCGCGCCGAGGTGCTCAAGAGGAAGTACGCCCG 392
QY 664 GCTGTGATGCTGGTCCATCGGCGTCATCACCTACATATGCTGTGTGGATACCCCGCG 723
Db 393 GAGGCCGACATCTGGAGCGTGGCGTCTGCTGCTACATCTTCTCGCGCGGCTGCTGCC 452

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:07:56 ; Search time 16 Seconds
(without alignments)
875.332 Million cell updates/sec

Title: us-09-960-643-2
Perfect score: 2513
Sequence: 1 MGRKEEDDCSWKKQTTNIR.....VKAGSSHCRCAGTGVCLIM 476

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	1152.5	45.9	370	2	US-08-878-989-19		Sequence 19, Appl
2	1152.5	45.9	370	4	US-09-272-796-19		Sequence 19, Appl
3	1152.5	45.9	370	4	US-09-457-0408-31		Sequence 31, Appl
4	990	39.4	343	2	US-08-878-989-5		Sequence 5, Appli
5	990	39.4	343	4	US-09-272-796-5		Sequence 5, Appli
6	747.5	29.7	501	4	US-09-734-030-2		Sequence 2, Appli
7	650.5	25.9	424	2	US-08-715-568A-1		Sequence 1, Appli
8	625	24.9	264	2	US-07-857-224B-24		Sequence 24, Appl
9	592	23.6	556	4	US-09-800-960-4		Sequence 4, Appli
10	587.5	23.4	264	2	US-07-857-224B-18		Sequence 18, Appl
11	578.5	23.0	543	4	US-09-529-093A-2		Sequence 2, Appli
12	571.5	22.7	565	4	US-09-800-960-2		Sequence 2, Appli
13	564	22.4	264	2	US-07-857-224B-19		Sequence 19, Appl
14	563.5	22.4	295	1	US-07-951-715A-23		Sequence 23, Appl
15	563.5	22.4	295	2	US-08-459-448A-23		Sequence 23, Appl
16	563.5	22.4	295	3	US-08-459-595A-23		Sequence 23, Appl
17	563.5	22.4	295	3	US-08-459-504B-23		Sequence 23, Appl
18	563.5	22.4	295	3	US-08-459-444-23		Sequence 23, Appl
19	563.5	22.4	295	4	US-09-547-422-23		Sequence 23, Appl
20	558	22.2	776	1	US-08-198-446B-17		Sequence 17, Appl
21	558	22.2	776	2	US-08-870-693-17		Sequence 17, Appl
22	558	22.2	821	1	US-08-198-446B-6		Sequence 6, Appli
23	558	22.2	821	2	US-08-870-693-6		Sequence 6, Appli
24	542.5	21.6	463	1	US-07-951-715A-25		Sequence 25, Appl
25	542.5	21.6	463	2	US-08-459-448A-25		Sequence 25, Appl
26	542.5	21.6	463	3	US-08-459-595A-25		Sequence 25, Appl
27	542.5	21.6	463	3	US-08-459-504B-25		Sequence 25, Appl

Sequence 0, Appli
Sequence 0, Appli
Sequence 17, Appl
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 20, Appl
Sequence 1, Appli
Sequence 1, Appli
Sequence 10, Appl
Sequence 2, Appli
Sequence 2, Appli
Sequence 24, Appl
Sequence 21, Appl
Sequence 23, Appl
Sequence 18, Appl
Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-878-989-19
; Sequence 19, Application US/08878989
; Patent No. 5885803
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,989
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank

; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT06
; CLONE: 827431
; US-09-272-796-5

Query Match 39.4%; Score 990; DB 4; Length 343;
Best Local Similarity 59.8%; Pred. No. 2.9e-79;
Matches 193; Conservative 49; Mismatches 73; Indels 8; Gaps 2;

QY 13 KQTNTIRKTFIFMEVLGSGAFSEVFLVQRLTGKLFALKCI-KKSPAFRDSLENEIAV 71
Db 5 KKHTDISSVYEIRERLGGSGAFSEVLAQERGSAAHLVALKCI PKKALRGKEALVENEIAV 64

QY 72 LKIKIHENIVTLDEIYESTHYLVNQLVSGGELFDRILRGVYVEKDALSLVIOQVLSAV 131
Db 65 LRRISPNIVALEDVHESPESHLYLAMELVGTGGELFDRIMERGSYTEKDALSLVIOQVLSAV 124

QY 132 KYLHENGIVHRDLKPNLLYLTPPEENSKIMITDFGLSKMEQNGIMSTACGTGCVYVAPEVL 191
Db 125 SYLHSGIVHRDLKPNLLYATPEFESKIMVSDFGLSKIQAQNLGTACGTGCVYVAPELL 184

QY 192 AQPKYKAVDCWISGIVITVILGCGPPFYEEETESKLFKEKIGYEFESFPFDDISESAK 251
Db 185 EQPKYKAVDVWALGVISVILGCGPPFYDESDPELFSQILRASVYEFDPFDDISESGK 244

QY 252 DFTCHLLEKDPNRYTCERKALSHPWIDGNLTALHRDIYPSVLSIQKNAFKSKWRQAFNA 311
Db 245 DFTIRHLERDLQKRTCCQALRDLWIFDWTGFGDRILGFVSEQIRKNFAWTHWKRAFNA 304

QY 312 AVVHHMRKLMNLSHPCVRPEVE 334
Db 305 LFLRHIRKL-----GQIPEGE 320

RESULT 6
US-09-734-030-2
; Sequence 2, Application US/09734030
; Patent No. 6461846
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M.
; APPLICANT: MERKLOV, Gennady
; APPLICANT: KETCHUM, Karen A.
; APPLICANT: WEI, Ming-Hui
; APPLICANT: DIFRANCESCO, Valentina
; APPLICANT: YAN, Chunhua
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000612

; CURRENT APPLICATION NUMBER: US/09/734,030
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/207,281
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: PRT
; ORGANISM: HUMAN
; US-09-734-030-2

Query Match 29.7%; Score 747.5; DB 4; Length 501;
Best Local Similarity 35.3%; Pred. No. 1.2e-57;
Matches 173; Conservative 86; Mismatches 170; Indels 61; Gaps 12;

QY 15 QTTNIRKTFIFMEVLGSGAFSEVFLVQRLTGKLFALKCIKKSAPFRD-----SSLENEI 69
Db 16 QPSEVTDYDLGQVIKTEEFCEIFRAKDKTKGLHTCKKFKQ-----RDGRKVRKAANEI 71

QY 70 AVLKKIHENIVTLDEIYESTHYLVNQLVSGGELFDRILRGVYVEKDALSLVIOQVLS 129
Db 72 GILKMKVHPNQLQVDFVTRKEYFIFLELATGREVFDWILDQGYSERDTSNVVRQVLE 131

QY 130 AVKYLHENGIVHRDLKPNLLYLTPPEENSKIMITDFGLSKMEQNGIMSTACGTGCVYVAPE 189
Db 132 AVAYLSLKIIVHRNLKLENLVYNRLKNSKIVISDFHLAKLE-NGLIKEPCGTPPEYLAPE 190

QY 190 VLAQPKYKAVDCWISGIVITVILGCGPPFYEEETESKLFKEKIGYEFESFPFDDISESAK 241
Db 191 VVGRQYRGVDCWISGIVITVILGCGPPFYEEETESKLFKEKIGYEFESFPFDDISESAK 250

QY 242 FMDDISESAKDFICHLEKDPNRYTCERKALSHPWIDGNLTALHRDIYPSVLSIQKNAFK 301
Db 251 YWDDISQAQKDLVTRLMEVEQDQRTAEAEISHEWISGNAASDKNIKDGVCQAIEKNFAR 310

QY 302 SKWRQAFNAAVVHHMRKLMNLSHPCVRPEVE-ETSRP-----SSPE 352
Db 311 AKWKAVRVTLMKRLR-----APEQSSATAAQSASATDITATPGAGGATATAA 359

QY 353 ITITEAPVLDHSHVALPALTLQPCQHRRTAPGGRSLNCLVNGSL--HISSSLVPMHQGS 410
Db 360 SGATSAPEGDAARAASDNVAPADRSATPATDG--SATPATDGSVTPATDGSITPATDGS 417

QY 411 L-----AAGPCGCCSSCLNIGSKGKSYCSEPTLLKANKKQNFSEVMV---PVRASG 461
Db 418 VTPVTVDRSATP-----ATDGRATPATPEESTVPTTQSSAMLATKATAATPEPAMAQP 467

QY 462 SSHCRAGOTG 471
Db 468 DSTAPEGATG 477

RESULT 7
US-08-715-568A-1
; Sequence 1, Application US/08715568A
; Patent No. 5856463
; GENERAL INFORMATION:
; APPLICANT: Prydz, Hans Peter Blankenborg
; APPLICANT: Brede, Gaute
; TITLE OF INVENTION: PSKH-1 Ribozymes and Uses in Disease
; TITLE OF INVENTION: Treatment
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik
; STREET: 600 South Avenue West
; CITY: Westfield
; STATE: NJ
; COUNTRY: USA
; ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/715,568A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Foley, Shawn P.
REFERENCE/DOCKET NUMBER: FORSK 3.0-002
TELEPHONE: 908-654-5000
TELEFAX: 908-654-7866
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-715-568A-1

Query Match 25.9%; Score 650.5; DB 2; Length 424;
Best Local Similarity 40.2%; Pred. No. 3.5e-49;
Matches 131; Conservative 62; Mismatches 98; Indels 35; Gaps 4;
QY 27 EVLGSAPFVFLVKQRLTGKLFALKCIKKSPAFDRDSSLENEIAVLKIKHENIVTLEDI 86
Db 102 EIIIGSGFSRVVRVHRATRPVAKMIETKYREGREVCESELVLRVRHANIQLVEV 161
QY 87 YESTHYIYVQLVSGGELFDRLERGVYTEKDALSVIQVLSAVKYLHENGIVHRDLKP 146
Db 162 FETQERYVMVMEATGGELFDRIIAKGSFETERDATRVLMQVLDGVRYLHALGITHRDLKP 221
QY 147 ENLLYLPPENSKIMITDGLSKMEONG---IMSTACGTPGYVAPVLAQKPYSKAVDCW 203
Db 222 ENLLYHPGTDSKIIITDFGLASARKKGGDCLMKTTCTGPTPIAPEVLVRKPYTNSVDMW 281
QY 204 SIGVITYILLGCGYPPFEETESKLFKEIKEGYEFESPFWDIISAKDFICHLEKDPN 263
Db 282 ALGVIAVLLSGTMTPEDDNRTLRQLRGKYSYSGEPWPSVSNLAKDFIDRLLTVDPG 341
QY 264 ERYTCEKALSHPWIDNTALHRDIYPSVLIQTKNFAKSKRQAFNAAVVHMKRLMN 323
Db 342 ARMTALQALRHPWV-----VSM-----AASSSMKNLHRS 370
QY 324 LHPGVRPEVENRPPETOASERSPS 349
Db 371 I-SQNLKRRASRCQSTKSAQSTRSS 395

RESULT 8
US-07-857-224B-24
Sequence 24, Application US/07857224B
Patent No. 5958784
GENERAL INFORMATION:
APPLICANT: Benner, Steven A.
TITLE OF INVENTION: Predicting Folded Structures of Proteins
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven A. Benner
STREET: Hadlaubstrasse 151
CITY: Zurich
STATE: none
COUNTRY: Switzerland
ZIP: (note: this is an international post code) CH-8052
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/857,224B
FILING DATE: 03/25/92
CLASSIFICATION: 436
PRIOR APPLICATION DATA: none

TELECOMMUNICATION INFORMATION:
TELEPHONE: (International) 41 1 632 2830
TELEFAX: (International) 41 1 262 2437
TELEX: none
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 264
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: protein
ORIGINAL SOURCE:
ORGANISM: human
FEATURE: Protein kinase; Table 8 Column 26
PUBLICATION INFORMATION:
AUTHORS:
AUTHORS: Hanks, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988
US-07-857-224B-24

Query Match 24.9%; Score 625; DB 2; Length 264;
Best Local Similarity 46.9%; Pred. No. 3e-47;
Matches 119; Conservative 50; Mismatches 81; Indels 4; Gaps 2;
QY 28 VLGSAPFVFLVKQRLTGKLFALKCIKKSPAFDRDSSLENEIAVLKIKHENIVTLEDIY 87
Db 8 LIGRGSFSRVVRVHRATRPVAKMIETKYREGREVCESELVLRVRHANIQLVEVF 67
QY 88 ESTHYIYVQLVSGGELFDRLERGVYTEKDALSVIQVLSAVKYLHENGIVHRDLKP 146
Db 68 ETOERYVMVMEATGGELFDRIIAKGSFETERDATRVLMQVLDGVRYLHALGITHRDLKP 127
QY 147 ENLLYLPPENSKIMITDGLSKMEONG---IMSTACGTPGYVAPVLAQKPYSKAVDCW 203
Db 128 ENLLYHPGTDSKIIITDFGLASARKKGGDCLMKTTCTGPTPIAPEVLVRKPYTNSVDMW 187
QY 204 SIGVITYILLGCGYPPFEETESKLFKEIKEGYEFESPFWDIISAKDFICHLEKDPN 263
Db 188 ALGVIAVLLSGTMTPEDDNRTLRQLRGKYSYSGEPWPSVSNLAKDFIDRLLTVDPG 247
QY 264 ERYTCEKALSHPW 277
Db 248 ARMTALQALRHPWV 261

RESULT 9
US-09-800-960-4
Sequence 4, Application US/09800960
Patent No. 6387677
GENERAL INFORMATION:
APPLICANT: YE, Jane et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001158
CURRENT APPLICATION NUMBER: US/09/800,960
CURRENT FILING DATE: 2001-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 556
TYPE: PRT
ORGANISM: Human
US-09-800-960-4
Query Match 23.6%; Score 592; DB 4; Length 556;
Best Local Similarity 33.2%; Pred. No. 7.5e-44;

Db 264 ADPAL--NVETEIEILKLNHPICIIKKNFFDA-EDYIVVLEMEGGELFDKVVGNKRL 319

QY 116 TEKDASLVVQVLSAVKYVHENGIVHRDLKPNLNYLTPPENSKIMITDFGLSK-MEONG 174

Db 320 KEATCKLYFYOMLLAVQYLHENGIVHRDLKPNLVLSQEECLIKITDFGHSKILGETS 379

QY 175 IMSTAGCTGYVAPEVLAQ--KPYSKAVDCWSIGVITVILLCGYPPPEYE-ETESKLPK 230

Db 380 LMTLCTGYTTLAPEVLSVGTAGYNRAVDCWSLGVITVILLCGYPPPESEHRTQVSLKDQ 439

QY 231 IKEGYEFESPFWDIDISESAKDFICHLLEKDPNERTCCKALSHPMI 277

Db 440 ITSGKYNFPEVWAEYSERALDVKKLLVVDPKARETTEALRHPWL 486

RESULT 12

US-09-800-960-2

; Sequence 2, Application US/098000960

; Patent No. 6387677

; GENERAL INFORMATION:

; APPLICANT: YE, Jane et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001158

; CURRENT APPLICATION NUMBER: US/09/800,960

; CURRENT FILING DATE: 2001-03-08

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 565

; TYPE: PRT

; ORGANISM: Human

US-09-800-960-2

Query Match 22.4%; Score 571.5; DB 4; Length 565;

Best Local Similarity 34.4%; Pred. No. 4.9e-42;

Matches 142; Conservative 69; Mismatches 151; Indels 51; Gaps 10;

QY 17 TNRKTFIFMEVLSGAFSEVFLVKQRLTGKLFALKCI--KKSPAFRDSSENEIAVLKK 74

Db 8 TRFTDDYQLFEELGKGFVSVRCVKKTSTQYAAKIIINTKLSARDHQKLERARICRL 67

QY 75 IKHENIVTLEDIYESTHYVLMOLVSGGELFDRLILRGVYTEKDASLVVQVLSAVKYL 134

Db 68 LKPNVIRLHDSISEGPHVLDVLTGGELFEDIVAREYVSEADASHCHIHQLESVNH 127

QY 135 HENGIVHRDLKPNLNYLTPPENSKIMITDFGLS---KMEQNGIMSTAGTGYVAPEVL 191

Db 128 HQHDIVHRDLKPNLLASKCKGAAVKLADFGLAIEVQGEQQAQWFGFA-GTGYLSPEVL 186

QY 192 AQPKYSKAVDCWSIGVITVILLCGYPPPEYEETESKLPKIKEGYEFESPFWDIDISESAK 251

Db 187 RKDPYKPDVIMACGVILVILVGPFPFDEQHLKYOQIKAGAYDFPSPEWDTVTPEAK 246

QY 252 DFICHLLEKDPNERTCCKALSHPMIDGNTA---LHRDIYPS-----V 291

Db 247 NLINQMLTINPAKRITADQALKHPWCQSRSTVASMHRQETVECLRKFNARKKLKAILT 306

QY 292 SLOIQNFAKSKWRQ-----AFNAAAVVHMRKLHNLHNSPGVRPEVENRPETQASETS 346

Db 307 TMLVSRNFSVGRSSAPASPAASAGLAGAAKSLINKSSDG---GVKKR-----KSSSV 359

QY 347 RPSSPETI-----TEAPVLDHVALPALTOPCQHGRRPTAPGGRS 388

Db 360 HLMPEQTIVVHNATDGIKSTES--CNTTDEDLKAAPLRTGNGSSVPEGRS 410

RESULT 13

US-07-857-224B-19

; Sequence 19, Application US/07857224B

; Patent No. 5958784

; GENERAL INFORMATION:

; APPLICANT: Benner, Steven A.

; TITLE OF INVENTION: Predicting Folded Structures of Proteins

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Steven A. Benner

; STREET: Hadlaubstrasse 151

; CITY: Zurich

; STATE: none

; COUNTRY: Switzerland

; ZIP: (note: this is an international post code) CH-8092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: Macintosh 7.0

; SOFTWARE: Microsoft Word

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/857,224B

; FILING DATE: 03/25/92

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA: none

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (International) 41 1 632 2830

; TELEFAX: (International) 41 1 262 2437

; TELEX: none

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 264

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE:

; DESCRIPTION: protein

; ORIGINAL SOURCE:

; ORGANISM: rat

; FEATURE: Protein kinase; Table 8 Column 21

; PUBLICATION INFORMATION:

; AUTHORS: Hanks, S. K.

; AUTHORS: Quinn, A. M.

; AUTHORS: Hunter, T.

; TITLE: The protein kinase family

; JOURNAL: Science

; VOLUME: 241

; PAGES: 42-52

; DATE: 1988

US-07-857-224B-19

Query Match 22.4%; Score 564; DB 2; Length 264;

Best Local Similarity 45.7%; Pred. No. 7.3e-42;

Matches 117; Conservative 42; Mismatches 91; Indels 6; Gaps 3;

QY 27 EVLGSAGFSEVFLVKQRLTGKLFALKCI--KKSPAFRDSSENEIAVLKKIKHENIVTLE 84

Db 7 EDIGKGAFFSVVRCVCKCTGHEYAAKIIINTKLSARDHQKLERARICRLKHSNIVRLH 66

QY 85 DIYESTTHYVLMOLVSGGELFDRLILRGVYTEKDASLVVQVLSAVKYVHENGIVHRDL 144

Db 67 DSISEGPHVLDVLTGGELFEDIVAREYVSEADASHCHIQILEAVLHCHQMGVVRDL 126

QY 145 KPNLNYLTPPENSKIMITDFGLSKMEQNGIMST---ACGTPGYVAPEVLAQPKYSKAVD 201

Db 127 KPNLNLASKLKGAAVKLADFGLA-IEVQGDQQAQWFGAGTGGYLSPEVLRKEAYGKPD 185

QY 202 CWSIGVITVILLCGYPPPEYEETESKLPKIKEGYEFESPFWDIDISESAKDFICHLLEK 261

Db 186 IWACGVILVILVGPFPFDEQHLKYOQIKAGAYDFPSPEWDTVTPEAKNLINQMLTIN 245

QY 262 PNERTCCKALSHPMI 277

Db 246 PAKRITAHEALKHPW 261

RESULT 14

US-07-951-715A-23

	Query Match	22.4%	Score 563.5;	DB 1;	Length 295;
	Best Local Similarity	41.8%	Pred. No. 9.5e-42;		
	Matches 120;	Conservative 56;	Mismatches 104;	Indels 7;	Gaps 4;
QY	27	EVLGSCASEFLVQRLTGKLFALKCI--KKSPAPEDSSLENEIAVLKKIKHENIVYLE	84		
		: : : :			
Db	5	EELGAGFSWRRVCYKTKSTQEYAAKLIINTKLSARDHOKLEAREARICLLKHPNIVRUH	64		
		: : : :			
QY	85	DIVESTHYLLVMQLVSGGELFDRLRGVVYTERDASLVITQQVLSAVKYLHENGIVHRDL	144		
		: : : :			

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..295
; OTHER INFORMATION: /note= "rat protein kinase II
; OTHER INFORMATION: protein sequence as shown in Figure 32."
US-08-459-448A-23

Query Match      22.4%; Score 563.5; DB 2; Length 295;
Best Local Similarity 41.8%; Pred. No. 9.5e-42;
Matches 120; Conservative 56; Mismatches 104; Indels 7; Gaps 4;

QY 27 EVLGSGAFSEVFLVKQRLTKLFALKCI--KSPAPFRDSSLENEIAVLKKIKHENIVTLE 84
Db 5 EELGKGAFSVVRCVKKTSTQEYAAKIINTKKLSARDHQKLEREARICRLKHPNIVRLH 64

QY 85 DIYESTHYLYVMQLVSGGELFDRILERGVYTEKDASLVIOQVLSAVKYLHENGIVHRDL 144
Db 65 DSISEEGFHYLVFDLVTGGELFEDIVAREYISEADASHCIHQLESVNIHQHDIYHRDL 124

QY 145 KPNLLVLTPEENSKIMITDEGLS--KMEQNGIMSTAGTGCYVAPEVLAQKPYSKAVD 201
Db 125 KPNLLASCKGAAYKLADFGLAIEVQGEQQAWFGFA-GTPLYSPVLRKDPYGPVPD 183

QY 202 CWSIGVITYILLGYPFPYEETESKLFEKTEGYEFESPFWDIDISEAKDFICHILLEKD 261
Db 184 IWACGVILYILLVGYPPFWEQDKLYQQIKAGAYDFPSPWDTVTPPEAKNLINQMLTIN 243

QY 262 PNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKNF-AKSKWROA 307
Db 244 PAKRITADQALKHPWVCQRSTVASMHRQETVECLRKFNARRKLKGA 290
```

Search completed: March 14, 2003, 14:10:30
Job time : 19 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:06:52 ; Search time 32 Seconds
(without alignments)
3064.951 Million cell updates/sec

Title: US-09-960-643-2
Perfect score: 2513
Sequence: 1 MKREEDCSWKKQTNR.....VKASGSHCRAGTGVCLIM 476

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2513	100.0	481	4 Q9V3J7	Q9V3J7 homo sapien
2	2509	99.8	476	4 Q96NX5	Q96NX5 homo sapien
3	2362	94.0	460	4 Q95523	Q95523 homo sapien
4	2327.5	92.6	477	11 Q91VB2	Q91VB2 mus musculus
5	1612	64.1	309	11 Q08763	Q08763 rattus norv
6	1246	49.6	357	4 Q9HD31	Q9HD31 homo sapien
7	1155.5	46.0	374	11 Q91VS8	Q91VS8 mus musculus
8	1070	42.6	343	11 Q70150	Q70150 rattus norv
9	1068	42.5	342	11 Q08767	Q08767 rattus norv
10	1064	42.3	343	11 Q9QYK9	Q9QYK9 mus musculus
11	1047	41.7	343	4 Q9NNY2	Q9NNY2 homo sapien
12	1008	40.1	348	5 Q9TXJ0	Q9TXJ0 caenorhabdi
13	1004	40.0	348	5 Q9UAH6	Q9UAH6 caenorhabdi
14	957.5	38.1	405	5 Q9V3I4	Q9V3I4 drosophila
15	747.5	29.7	501	4 Q8WTT8	Q8WTT8 homo sapien
16	747.5	29.7	501	4 Q9BQC9	Q9BQC9 homo sapien

17	739	29.4	512	11 Q8VD20	Q8VD20 mus musculus
18	736	29.3	504	11 Q63092	Q63092 rattus norv
19	703	28.0	421	13 Q9YGM4	Q9YGM4 fugu rubrip
20	692.5	27.6	502	11 Q63892	Q63892 rattus sp.
21	674	26.8	385	13 Q98TW2	Q98TW2 xenopus lae
22	651.5	25.9	349	5 Q8SSR6	Q8SSR6 dictyosteli
23	643.5	25.6	424	11 Q91YA2	Q91YA2 mus musculus
24	643	25.6	373	3 Q9HF33	Q9HF33 arthrobotry
25	635	25.3	420	3 Q42627	Q42627 collettotric
26	631.5	25.1	415	3 Q96UI7	Q96UI7 neurospora
27	628.5	25.0	473	4 Q9H0Q5	Q9H0Q5 homo sapien
28	623	24.8	404	3 Q9Y899	Q9Y899 emeritella
29	621	24.7	513	5 Q15865	Q15865 plasmodium
30	615.5	24.5	533	5 Q9NH59	Q9NH59 caenorhabdi
31	615.5	24.5	571	5 Q9NH58	Q9NH58 caenorhabdi
32	610	24.3	720	5 Q62305	Q62305 caenorhabdi
33	608	24.2	518	5 Q9NH57	Q9NH57 caenorhabdi
34	606.5	24.1	385	4 Q96QS6	Q96QS6 homo sapien
35	605.5	24.1	433	11 Q9JLM6	Q9JLM6 mus musculus
36	604.5	24.1	589	11 Q63094	Q63094 rattus norv
37	603	24.0	350	5 Q9NG91	Q9NG91 caenorhabdi
38	603	24.0	482	5 Q9U6Q0	Q9U6Q0 caenorhabdi
39	603	24.0	520	5 Q9NH60	Q9NH60 caenorhabdi
40	602	24.0	478	13 Q9YHB8	Q9YHB8 gallus gall
41	602	24.0	527	4 Q9Y2H4	Q9Y2H4 homo sapien
42	600	23.9	530	5 Q00168	Q00168 drosophila
43	599.5	23.9	493	5 Q9V495	Q9V495 drosophila
44	598.5	23.8	489	13 Q93559	Q93559 gallus gall
45	598.5	23.8	499	6 Q95266	Q95266 sus scrofa

ALIGNMENTS

RESULT 1

Q9V3J7 Q9V3J7 PRELIMINARY; PRT; 481 AA.
AC Q9V3J7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 53.6 kDa protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rhodes S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL049688; CAB41259.1; .
DR HSSP; Q63450; IA06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
DR PROSITE; PS00107; PROTEIN_KINASE_APP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_SP; 1.
KW Hypothetical protein; Serine/threonine-protein kinase.
FT NON_TER 1
SQ SEQUENCE 481 AA; 53588 MW; 3C5A4F9B6765333E CRC64;

Query Match 100.0%; Score 2513; DB 4; Length 481;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRKEEDDCSSWKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAF 60
|||||
Db 6 MGRKEEDDCSSWKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAF 65
|||||
QY 61 RDSSLENEIAVLKKIKHENIVTLEDIYESTHYILVMQVLVSGGELFDRILRGVYTEKDA 120
|||||
Db 66 RDSSLENEIAVLKKIKHENIVTLEDIYESTHYILVMQVLVSGGELFDRILRGVYTEKDA 125
|||||
QY 121 SLVIOQVLSAVKYLHENGIVHRLDKPENLLYLTPEENSKIMITDGLSKMQNGIMSTAC 180
|||||
Db 126 SLVIOQVLSAVKYLHENGIVHRLDKPENLLYLTPEENSKIMITDGLSKMQNGIMSTAC 185
|||||
QY 181 GTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFES 240
|||||
Db 186 GTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFES 245
|||||
QY 241 PFWDIISAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVLSIQKNFA 300
|||||
Db 246 PFWDIISAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVLSIQKNFA 305
|||||
QY 301 KSKWROAFNAAAVVHHMRKLMNLHSPGVPRPEVENRPPETOASSETSRPSSPEITITEAPV 360
|||||
Db 306 KSKWROAFNAAAVVHHMRKLMNLHSPGVPRPEVENRPPETOASSETSRPSSPEITITEAPV 365
|||||
QY 361 LDHSVALPALTOQLPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHOGSLAAGPCGCCS 420
|||||
Db 366 LDHSVALPALTOQLPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHOGSLAAGPCGCCS 425
|||||
QY 421 SCNLGSKGKSSYCSPTLLKANKKQNFSEVMPVKASGSSHCRAGOTGVCLIM 476
|||||
Db 426 SCNLGSKGKSSYCSPTLLKANKKQNFSEVMPVKASGSSHCRAGOTGVCLIM 481
|||||

RESULT 2

Q96NX5 PRELIMINARY; PRT; 476 AA.
ID Q96NX5; AC Q95523; PRELIMINARY; PRT; 460 AA.
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Calcium/calmodulin-dependent protein kinase I gamma.
GN CAMK1G.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
-OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20113118; PubMed=10645953;
RA Schutte B.C., Bjork B.C., Coppage K.B., Malik M.I., Gregory S.G.,
RA Scott D.J., Brentzell L.M., Watanabe Y., Dixon M.J., Murray J.C.;
RT "A preliminary gene map for the Van der Woude syndrome critical region
RT derived from 900 kb of genomic sequence at 1q32-q41.";
RL Genome Res. 10:81-94(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Bjork B.C., Watanabe Y., Murray J.C., Schutte B.C.;
RT "Characterization of the human ortholog of rat Cam Kinase I gamma
RT (CamKig) at 1q32-q41.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF426261; AAL28100.1; -
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001230; Prenyl_site.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS00294; PHENYLATION; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.

SQ SEQUENCE 476 AA; 53128 MW; DD0AEAA016E7506E CRC64;
Query Match 99.8%; Score 2509; DB 4; Length 476;
Best Local Similarity 99.6%; Pred. No. 4.8e-190;
Matches 474; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRKEEDDCSSWKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAF 60
|||||
Db 1 MGRKEEDDCSSWKQTTNIRKTFIFMEVLGSGAFSEVFLVKQRLTGKLFALKCIKKSPAF 60
|||||
QY 61 RDSSLENEIAVLKKIKHENIVTLEDIYESTHYILVMQVLVSGGELFDRILRGVYTEKDA 120
|||||
Db 61 RDSSLENEIAVLKKIKHENIVTLEDIYESTHYILVMQVLVSGGELFDRILRGVYTEKDA 120
|||||
QY 121 SLVIOQVLSAVKYLHENGIVHRLDKPENLLYLTPEENSKIMITDGLSKMQNGIMSTAC 180
|||||
Db 121 SLVIOQVLSAVKYLHENGIVHRLDKPENLLYLTPEENSKIMITDGLSKMQNGIMSTAC 180
|||||
QY 181 GTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFES 240
|||||
Db 181 GTPGYVAPEVLAQKPYSKAVDCWSIGVITYILLCGYPFPYEETESKLFKEIKEGYEFES 240
|||||
QY 241 PFWDIISAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVLSIQKNFA 300
|||||
Db 241 PFWDIISAKDFICHLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVLSIQKNFA 300
|||||
QY 301 KSKWROAFNAAAVVHHMRKLMNLHSPGVPRPEVENRPPETOASSETSRPSSPEITITEAPV 360
|||||
Db 301 KSKWROAFNAAAVVHHMRKLMNLHSPGVPRPEVENRPPETOASSETSRPSSPEITITEAPV 360
|||||
QY 361 LDHSVALPALTOQLPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHOGSLAAGPCGCCS 420
|||||
Db 361 LDHSVALPALTOQLPCQHGRRPTAPGGRSLNCLVNGSLHSSSLVPMHOGSLAAGPCGCCS 420
|||||
QY 421 SCNLGSKGKSSYCSPTLLKANKKQNFSEVMPVKASGSSHCRAGOTGVCLIM 476
|||||
Db 421 SCNLGSKGKSSYCSPTLLKANKKQNFSEVMPVKASGSSHCRAGOTGVCLIM 476
|||||
RESULT 3
Q95523 PRELIMINARY; PRT; 460 AA.
ID Q95523; AC Q95523; PRELIMINARY; PRT; 460 AA.
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE DJ272116.1 (Rat Ca2+/calmodulin dependent protein kinase like
DE protein).
DE DJ272116.1.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
-OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AL023754; CAAL9296.1; -
DR HSSP: Q63450; 1A06.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 460 AA; 51486 MW; A16971DC50344496 CRC64;
Query Match 94.0%; Score 2362; DB 4; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.9e-178;

DR	InterPro; IPR001245; Tyr_kinase.
DR	Pfam; PF00069; kinase; 1.
DR	ProDom; PD000001; Euk_pkinase; 1.
DR	SMART; SM00219; TyrKc; 1.
DR	PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_St; UNKNOWN_1.
DR	ATP-binding; Kinase; Transferase; hypothetical protein.
KQ	SEQUENCE 477 AA; 53296 MW; 3A61rBDC49A7BB56 CRC64;
SQ	
	Query Match - 92.6%; Score 2327.5; DB 11; Length 477;
	Best Local Similarity 92.0%; Pred. No. 1.1e-175;
	Matches 439; Conservative 14; Mismatches 23; Indels 1; Gaps 1;
QY	1 MGRKEEDCSSWKQTNIKTFTFMEVLGSGAFSEVFLVKQRILTKGLFKALCKIKSPAF 60
DB	1 MGRKEEDCSSWKQTNIKTFTFMEVLGSGAFSEVFLVKQRILTKGLFKALCKIKSPAF 60
QY	61 ROSSLENAVLAKRKIHENIVTLIEDIYSTHYVLMQLVSGGELFDRLERGVTERKA 120
DB	61 ROSSLENAVLAKRKIHENIVTLIEDIYSTHYVLMQLVSGGELFDRLERGVTERKA 120
QY	121 SLVIQOVLSAVKYLHENGIVHRDLAKPENLLYLTPEENSKIMITDFGLSKMQNGINSTAC 180
DB	121 SLVIQOVLSAVKYLHENGIVHRDLAKPENLLYLTPEENSKIMITDFGLSKMQNGINSTAC 180
QY	181 GTPGYVAPEVLQAQPKYSKAVDCWSIGVTYITLLCGYPFPYEETESKLFKIKEGYEYES 240
DB	181 GTPGYVAPEVLQAQPKYSKAVDCWSIGVTYITLLCGYPFPYEETESKLFKIKEGYEYES 240
QY	241 PFWDIDISAKDFICHLLEKOPNERYTCEKALSHLPWIDGNHTALHRDIYPSVSIQKNFA 300

Qy	301	KSXWROAFNAAVVHHMKRLHNLHSPGVREVENRPPETQASSETSRPSPSEITTEAPV	360
Db	301	KSXWROAFNAAVVHHMKRLHNLHSPGVREVENRPPVPAPEVSRPDHSSITEADI	360
Qy	361	LDHSVALPALTOLPCQHQHRRTPAP-GRGSLNCLNGSLHISSSLVPMHQGSLAAGPCGCC	419
Db	361	LDPSITLPAITLPCSHSRSPAPSGGSLNCLNGSLRISSSLVPMQOGLATGPCGC	420
Qy	420	SSCLTNGSGKGSYSYCEPTPLLLKKANKKONFKSEVWPVKASGSSHCRAGQTGVCLIM	476
Db	421	SSCLTNGNKGKGSYSYCEPTPLFRKANKKONFKSEVWPVKAGSGTHCRGGQTGVCLYM	477
RESULT 5			
	008763	PRELIMINARY: PRT: 309 AA.	
ID	008763	PRELIMINARY: PRT: 309 AA.	
AC	008763;		
DT	01-JUL-1997 (TrEMBLrel. 04, Created)		
DT	01-JUL-1997 (TrEMBLrel. 04, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	Protein kinase (Fragment).		
DE	Rattus norvegicus (Rat)..		
OS	Rattus norvegicus (Rat)..		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RX	MEDLINE=97228532; PubMed=9074610;		
RT	Yokokura H., Terada O., Naito Y., Hidaka H.;		
RT	"Isolation and comparison of rat CDNA's encoding Ca2+/calmodulin-		
RT	dependent protein kinase I isoforms.";		
RL	Biochim. Biophys. Acta 1338:8-12(1997).		
CC	-I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.		
DR	EMBL; D86557; BAAL9880.1; -;		
DR	HSP; O63450; IAU0.		
DR	InterPro; IPR000719; Euk_pkinase.		
DR	InterPro; IPR002290; Ser_thr_pkinase.		

```
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW Serine/threonine-protein kinase.
FT NON_TER 309 309
SQ SEQUENCE 309 AA; 35623 MW; 9162487561CF44E7 CRC64;

Query Match 64.1%; Score 1612; DB 11; Length 309;
Best Local Similarity 98.4%; Pred. No. 2.2e-119;
Matches 304; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGRKEEDCCSWKKQTNIIRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCIKKSPAF 60
Db 1 MGRKEEDCCSWKKQTNIIRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCIKKSPAF 60
Qy 61 ROSSLENEIATVLKIKHENIVTLEDIYESTHYLYVMOLVSGGELFDRILRGVYTEKDA 120
Db 61 ROSSLENEIATVLKIKHENIVTLEDIYESTHYLYVMOLVSGGELFDRILRGVYTEKDA 120
Qy 121 SLVIOQVL SAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 180
Db 121 SLVIOQVL SAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNGIMSTAC 180
Qy 181 GTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKEGYEFES 240
Db 181 GTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKEGYEFES 240
Qy 241 PEWDIDISESAKDFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKNA 300
Db 241 PEWDIDISESAKDFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKNA 300
Qy 301 KSKWROAFN 309
Db 301 KSKWROAFN 309

RESULT 6
Q9HD31 ID Q9HD31 PRELIMINARY; PRT; 357 AA.
AC Q9HD31;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CamK1-like protein kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Verploegen S., Koenderman L., Coffey P.J.;
RT "Identification and characterization of CKLIK: a novel granulocyte
Ca2+/calmodulin-dependent kinase."
RL Blood 0:0-0(2000).
DR EMBL: AF286366; AAG00534.1; -.
DR HSP: Q63450; 1A06.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 357 AA; 40189 MW; 1FA184EEFA976FB4 CRC64;

Query Match 49.6%; Score 1246; DB 4; Length 357;
Best Local Similarity 73.1%; Pred. No. 2.3e-90;
Matches 237; Conservative 40; Mismatches 45; Indels 2; Gaps 2;

Qy 1 MGRKEEDCCSWKKQTNIIRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCI-KKSPA 59
Db 1 MARENGESSSSWKKQADIKFETKLTGTGAFSEVFLAEAEKATGKLFVAVKIPKALK 60
Qy 60 FROSSLENEIATVLKIKHENIVTLEDIYESTHYLYVMOLVSGGELFDRILRGVYTEK 119
Db 61 GKESSENEIATVLKIKHENIVALEDIYESPNIHLYVMOLVSGGELFDRIVEKGFTEK 120
Qy 120 ASLVIOQVL SAVKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNG-IMST 178
Db 121 ASTLIROVLDAVYVLRMGIVHRDLKPENLLYSQDEESKIMISDFGLSKMEGKGVNST 180
Qy 179 ACGTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGYPPFYETESKLFKEIKEGYEF 238
Db 181 ACGTPGYVAPEVLAQPKYSKAVDCWSIGVITYILLCGYPPFYDENSKLFEQILKAEYEF 240
Qy 239 ESPWDIDISESAKDFICHLLEKDPNERYTCEKALSHPWIDGNTALHRDIYPSVSLQIQKN 298
Db 241 DSPYWDIDISESAKDFINLMEKDPNKRYTCEQAAHFWIAGDTALNKNIHESVAQIRKN 300
Qy 299 FAKSKWROAFNAAAVVHMRKLHM 322
Db 301 FAKSKWROAFNATAVVRMRKLHL 324

RESULT 7
Q91YS8 ID Q91YS8 PRELIMINARY; PRT; 374 AA.
AC Q91YS8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to calcium/calmodulin-dependent protein kinase I.
GN A1505105.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014825; AAH14825.1; -.
DR MGI: 2141437; A1505105.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 374 AA; 41624 MW; 37889CDA717D3AB2 CRC64;

Query Match 46.0%; Score 1155.5; DB 11; Length 374;
Best Local Similarity 69.6%; Pred. No. 3.5e-83;
Matches 222; Conservative 43; Mismatches 51; Indels 3; Gaps 3;

Qy 12 WKQTNIIRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCI-KKSPAPRDSLENEIA 70
Db 10 W-KQAEIRDIYDFRDVLGTGAFSEVFLAEKRTQKLVAKIACIAKALEGEGSMENEIA 68
Qy 71 VLKTIKHENIVTLEDIYESTHYLYVMOLVSGGELFDRILRGVYTEKDAVLQVLSA 130
Db 69 VLHKIKHPNIVALDDIYEGGHLVLMOLVSGGELFDRIVEKGFTEKDAKRLFQVLDA 128
Qy 131 VKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNG-IMSTACGTFGYVAPE 189
Db 131 VKYLHENGIVHRDLKPENLLYLTPEENSKIMITDFGLSKMEQNG-IMSTACGTFGYVAPE 189
```


DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MCAMK1-BETA2 protein (pregnancy upregulated NONBIQUITOUS
DE CA2+/calmodulin-dependent kinase Pnck).
GN Pnck OR CAMK1B OR MCAMK1-BETA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueda T., Minami Y.;
RT "mcamk1-beta2";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
TX TISSUE=BRAIN;
RX MEDLINE=20139438; PubMed=10673339;
RA Gardner H.P., Rajan J.V., Ha S.I., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A., Marquis S.T., Chodosh L.A.;
RT "Cloning, Characterization, and Chromosomal Localization of Pnck, a
RT Ca2+/Calmodulin-Dependent Protein Kinase";
RL Genomics 63:279-288(2000).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB023027; BAA87926.1; -;
DR EMBL; AF181984; AAF29157.1; -;
DR HSSP; Q63450; 1A06.
DR MGD; MGI:1347357; Pnck.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 343 AA; 38519 MW; 1B4A28D367A936E CRC64;

Query Match 42.3%; Score 1064; DB 11; Length 343;
Best Local Similarity 59.2%; Pred. No. 5.4e-76;
Matches 206; Conservative 55; Mismatches 75; Indels 12; Gaps 3;

QY 13 KQQTNRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCI-KKSPAFRDSSLENEIAV 71
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 5 KQKTEISSVYEIRKLGSGAFSEVLAQERGAHLVAKCIPKALRGKALVENEIAV 64
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
QY 72 LKKIKHENIVTLEDIYESTHYLYVMQLVSGGELEFDRLERGVYTEKDALSLVIOQVLSAV 131
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 65 LRRISHPNVALEDVHESPSHLYLAMELVTGGELFDRLIMERGSYTEKDALSHLVGQVLGAV 124
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
QY 132 KYLHENGIVHRDLKPENLYLTPENSKIMITDFGLSKMEQNGIMSTACGTPGYVAPEVL 191
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 125 SYLHSLGIVHRDLKPENLYLATPFEDSKIMVSDFGLSKIQAGNMLGTACGTPGYVAPELL 184
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
QY 192 AQKPYSKAVDCWSIGVITVILLCGYPFYEETESKLEKIGEVFEFSPFDDISESAK 251
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 185 EQKPYKAVDVWALGVISVILLCGYPFYEETESKLEKIGEVFEFSPFDDISESAK 244
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
QY 252 DFICHLLEKDPNERTCEKALSHPWIDGNLTALHRDIYPSVLSIQKNFAKSKWROAFNAA 311
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 245 DFIRHLLEKDPKRTCCQALQHLWISGDAFDRDILGVSQSIQKNFARTHWKRAFNAT 304
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
QY 312 AVVHMRKLMHNLHSPGVPRPEVENRPPETQASRSPSPETITEAP 359
: | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 305 SFLRHIRKL-----GOSPEGE-----EASRCQMTRHSHPLGLTQSP 341
: | | | | | | | | : | | | | | | | | : | | | | | : | | | | |

RESULT 11
Q9NNY2 PRELIMINARY; PRT; 343 AA.
ID Q9NNY2

AC Q9NNY2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Ca2+/calmodulin-dependent protein kinase I.
GN CAMKI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97432815; PubMed=9286695;
RA Brenner V., Nyakatura G., Rosenthal A., Platzer M.;
RT "Genomic organization of two novel genes on human Xq28: compact head
RT to head arrangement of IDH gamma and TRAP delta is conserved in rat
RT and mouse";
RL Genomics 44:8-14(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Platzer M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U52111; AAF4509.1; -;
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 343 AA; 38500 MW; E954AE89DE5608C9 CRC64;

Query Match 41.7%; Score 1047; DB 4; Length 343;
Best Local Similarity 60.7%; Pred. No. 1.2e-74;
Matches 201; Conservative 52; Mismatches 70; Indels 8; Gaps 2;

QY 13 KQQTNRKTFIFMEVLGSGAFSEVFLVKORLTGKLFALKCI-KKSPAFRDSSLENEIAV 71
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 5 KQKTEISSVYEIRKLGSGAFSEVLAQERGAHLVAKCIPKALRGKALVENEIAV 64
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
QY 72 LKKIKHENIVTLEDIYESTHYLYVMQLVSGGELEFDRLERGVYTEKDALSLVIOQVLSAV 131
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 65 LRRISHPNVALEDVHESPSHLYLAMELVTGGELFDRLIMERGSYTEKDALSHLVGQVLGAV 124
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
QY 132 KYLHENGIVHRDLKPENLYLTPENSKIMITDFGLSKMEQNGIMSTACGTPGYVAPEVL 191
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 125 SYLHSLGIVHRDLKPENLYLATPFEDSKIMVSDFGLSKIQAGNMLGTACGTPGYVAPELL 184
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
QY 192 AQKPYSKAVDCWSIGVITVILLCGYPFYEETESKLEKIGEVFEFSPFDDISESAK 251
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 185 EQKPYKAVDVWALGVISVILLCGYPFYEETESKLEKIGEVFEFSPFDDISESAK 244
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
QY 252 DFICHLLEKDPNERTCEKALSHPWIDGNLTALHRDIYPSVLSIQKNFAKSKWROAFNAA 311
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 245 DFIRHLLEKDPKRTCCQALQHLWISGDAFDRDILGVSQSIQKNFARTHWKRAFNAT 304
||| : | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
QY 312 AVVHMRKLMHNLHSPGVPRPEVENRPPETQASRSPSPETITEAP 342
: | | | | | | | | : | | | | | | | | : | | | | | : | | | | |
Db 305 SFLRHIRKL-----GQIPGEGACSEQGM 328
: | | | | | | | | : | | | | | | | | : | | | | | : | | | | |

RESULT 12
Q9TXJ0 PRELIMINARY; PRT; 348 AA.
AC Q9TXJ0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

```
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 39.1 kDa protein.
GN K07A9.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Davidson S., O'Neal D.;
RT "The sequence of C. elegans cosmid K07A9.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF099924; AAF23187.1; -.
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;
KW Transferase.
SQ SEQUENCE 348 AA; 39124 MW; 88525C39080A709F CRC64;

Query Match 40.1%; Score 1008; DB 5; Length 348;
Best Local Similarity 57.3%; Pred. No. 1.5e-71;
Matches 185; Conservative 68; Mismatches 68; Indels 2; Gaps 2;

QY 19 IRKTFIPEVLGSGAFSEFLVKQRL-TGKLFALKCI-KKSPAFRDSLSLENEIAVLKLIK 76
Db 18 IREKYDFRDLGTGAFSKVFLAESKSDAGQMYAVKCIDKALKGKEESLENEIKVLRLK 77

QY 77 HENIVTLEDIYESTHYLYMQLVSGELFDRILRGVYTEKSDASLVIOQVLSAVKYLHE 136
Db 78 HNNIVQLFDIYDEKQFVYLMVLTGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 137 NGIVHRDLKPENLLYLTPEENSKIMTDFGLSKMEQNGIMSTACGTPGYVAPEVLAQPY 196
Db 138 NGVYHRDLKPENLLYNNQDESDKIMISDFGLSKTEDSGVMATACGTPGYVAPEVLAQPY 197

QY 197 SKAVDCWSIGVITVILLCGYPPEYTESKLFKEIKEGYEFESPFWDIDISAKDFICH 256
Db 198 GKAVDWSIGVIAYILLCGYPPEYDESANLFAQIIKGEYEDAPYWDQISGSAKDFITH 257

QY 257 LLEKDPNERYTCEKALSHPWIDGNLTALHRDIYPSVSLQIOKNAKSKWRQAFNAAVVHH 316
Db 258 LMCCDPEARFTCDALSHPWISGNTAYTHDHTGTVAVHLKSLAKRWNKKAYNAAAIAIQ 317

QY 317 MRKLHMLHSPGYRPEVENRPE 339
Db 318 LQMLRLSSNRLQKQASQOQPE 340

RESULT 13
Q9UAH6
```

```
ID Q9UAH6 PRELIMINARY; PRT; 348 AA.
AC Q9UAH6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Ca2+/calmodulin-dependent protein kinase I.
GN CMK-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99357789; PubMed=10428833;
RA Eto K., Takahashi N., Kimura Y., Masubo Y., Arai K., Muramatsu M.,
RA Tokumitsu H.;
RT "Ca2+/calmodulin-dependent protein kinase cascade in Caenorhabditis
RT elegans. Implication in transcriptional activation.";
RL J. Biol. Chem. 274:22556-22562(1999).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB021864; BAA82674.1; -.
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 348 AA; 39066 MW; 24A2633BF4D88DB0 CRC64;

Query Match 40.0%; Score 1004; DB 5; Length 348;
Best Local Similarity 57.3%; Pred. No. 3e-71;
Matches 185; Conservative 67; Mismatches 69; Indels 2; Gaps 2;

QY 19 IRKTFIPEVLGSGAFSEFLVKQRL-TGKLFALKCI-KKSPAFRDSLSLENEIAVLKLIK 76
Db 18 IREKYDFRDLGTGAFSKVFLAESKSDAGQMYAVKCIDKALKGKEESLENEIKVLRLK 77

QY 77 HENIVTLEDIYESTHYLYMQLVSGELFDRILRGVYTEKSDASLVIOQVLSAVKYLHE 136
Db 78 HNNIVQLFDIYDEKQFVYLMVLTGELFDRIVAKSGSYTEQDASNLIRQVLEAVGFMD 137

QY 137 NGIVHRDLKPENLLYLTPEENSKIMTDFGLSKMEQNGIMSTACGTPGYVAPEVLAQPY 196
Db 138 NGVYHRDLKPENLLYNNQDESDKIMISDFGLSKTEDSGVMATACGTPGYVAPEVLAQPY 197

QY 197 SKAVDCWSIGVITVILLCGYPPEYTESKLFKEIKEGYEFESPFWDIDISAKDFICH 256
Db 198 GKAVDWSIGVIAYILLCGYPPEYDESANLFAQIIKGEYEDAPYWDQISGSAKDFITH 257

QY 257 LLEKDPNERYTCEKALSHPWIDGNLTALHRDIYPSVSLQIOKNAKSKWRQAFNAAVVHH 316
Db 258 LMCCDPEARFTCDALSHPWISGNTAYTHDHTGTVAVHLKSLAKRWNKKAYNAAAIAIQ 317

QY 317 MRKLHMLHSPGYRPEVENRPE 339
Db 318 LQMLRLSSNRLQKQASQOQPE 340

RESULT 14
Q9V314 PRELIMINARY; PRT; 405 AA.
ID Q9V314;
AC Q9V314; Q9V4D3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CamKI protein.
GN CAMKI OR CG1495.
OS Drosophila melanogaster (Fruit fly).
```


OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephytridae; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
OX	[1]
RP	SEQUENCE FROM N.A.
RP	TISSUE=RETINA.
RC	MEDLINE=99030403; PubMed=9813038;
RC	Xu X.Z.S., Wes P.D., Chen H., Li H.S., Yu M., Morgan S., Liu Y.,
RA	Montell C.;
RA	"Retinal targets for calmodulin include proteins implicated in
RT	synaptic transmission.";
RT	J. Biol. Chem. 273:31297-31297(1998).
RL	[2]
RP	SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RP	STRAIN=BERKELEY;
RC	MEDLINE=20196006; PubMed=107311132;
RC	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA	Balleg R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J.J., Brokstein P., Brotter P.,
RA	Burtis K.C., Busam D.A., Butler H., Cardieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA	Hosdin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA	Jalali M., Kalush F., Karpén G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Moult S.D., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissensbach J.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT	"The genome sequence of Drosophila melanogaster.";
RL	Science 287:2185-2195(2000).
CC	- - ALTERNATIVE PRODUCTS; 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC	SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC	- - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC	EMBL: Y17917; CAA76937.1; -
CC	EMBL: AE003844; AAF59343.1; -
DR	EMBL: AE003844; AAF59344.1; -
DR	HSP: Q63450; 1A06.
DR	FlyBase: FBgn0016126; CamK1.
DR	InterPro: IPR000719; Euk_pkinase.
DR	InterPro: IPR002290; Ser_thr_pkinase.
DR	Pfam: PF00069; pkinase; 1.
DR	ProDom: PD000001; Euk_pkinase; 1.
DR	SMART: SM00220; S_TKC_1
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR	KW ATP-binding; Alternative splicing; Kinase;
KW	Serine/threonine-protein kinase; Transferase.

FT	VARSPLIC	171	190	PENLLYVSPDDSKIMLSDF -> VRQFYLIYIFNDHLAL
FT				KGK (IN SHORT ISOFORM).
FT	VARSPLIC	191	405	MISSING (IN SHORT ISOFORM).
SEQ	SEQUENCE	405 AA;	451113 MW;	BD60C6C981CFBc4 CRC64;
Query Match 38.1%; Score 957.5; DB 5; Length 405;				
Best Local Similarity 55.4%; Pred. No. 1.8e-67;				
Matches 190; Conservative 64; Mismatches 70; Indels 19; Gaps				
QY	2 GRK-EEDDCSSWKQTTNRKTYFIMEVLGSGAFSEVFLVKQR-ITGKLFLAKCI-KKSP	58		
Db	10 GKRAKAKLKLNLKQ-VSTEEKYNLHGLLTGAFSEVRLAESKDSGPHFAVKIIDKAL	68		
QY	59 AFRODSLENEIAVLK-----KTKHENIVTLEDIVESTTHYLVLMQVSGG	103		
Db	69 KGKEESLENEIRVLRFSANHFQKCLNTRUTHNIVOLLLEYEDKSKVYLWMLVTGG	128		
QY	104 ELFDRLIRLGRVYTEKDALSLVIOQVLSAVKYLHENGIVHRDLKPENLLYITPEENSKIMIT	163		
Db	129 ELFDRIVERKGSYTEKDASHLRQILEAVDMHEQGVVHRDLKPENLLYVSPDDSKIMIS	188		
QY	164 DFLGSKMEQNGIMSTACGTPGVYAPVPLVLAQKPYSKAVDCWSIGVTYIILCGYPPFYET	223		
Db	189 DFLGSKMEDSGIMATACGTPGVYAPVPLVLAQKPYKAVDWSIGVTSYIILCGYPPFYDEN	248		
QY	224 ESKLFEKIKGYYEFESPEWDDISSEAKDFICHLLEKDPNERYTCEKALSHPWIDGNAL	283		
Db	249 DANLFAQLKGFDEFDPSPWDELSSEAKHIFIKLMCVTEYKRYTCKQALGHANWISGNEAS	308		
QY	284 HRDIYPSVSLQIQKNAFSAKSKWRQAFNAANAAVHHMRKLLHNLHS	326		
Db	309 SRNIHGTVTSQLKKNFAKSRWKQAYYAATVIRQMORMALNSNS	351		
RESULT 15				
Q8WTT8	IC	Q8WTT8	PRELIMINARY;	PRT; 501 AA.
AD	Q8WTT8;			
DT	01-MAR-2002 (TrEMBLrel. 20, Created)			
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Hypothetical 54.4 kDa protein.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	Strausberg R.;			
RL	Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.			
RL	[2]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE=LUNG;			
RA	Strausberg R.;			
RL	Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; BC019256; AAH19256.1; -			
DR	EMBL; BC017363; AAH17363.1; -			
DR	InterPro; IPR000104; Antifreeze_1.			
DR	InterPro; IPR000719; Euk_pkinase.			
DR	InterPro; IPR002290; Ser_thr_pkinase.			
DR	InterPro; IPR001245; Tyr_pkinase.			
DR	Pfam; PF00069; pkinase; 1.			
DR	PRINTS; PR00308; ANTIFREEZE1.			
DR	PRODOM; PD000001; Euk_pkinase; 1.			
DR	SMART; SM00220; S_TKC; 1.			
DR	SMART; SM00219; TyrKC; 1.			
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.			
KW	Hypothetical protein			
SEQ	SEQUENCE	501 AA;	1D6BF8BA0AD1D70	CRC64;
Query Match 29.7%; Score 747.5; DB 4; Length 501;				
Best Local Similarity 35.9%; Pred. No. 9.4e-51;				

		Matches 175; Conservative 81; Mismatches 176; Indels 55; Gaps 12;			
Qy	15	QTTNIRKTFIFMEVGLSGAFSEVFLVKQRLTGKLFALKCIKKSPAFRD-----SSLNEI	69		
Db	16	QPSVTDYDLGQVITKEEPEIFRAKDKTTGKLHTCKKFKQK----RDGRKVRKAAKNEI	71		
Qy	70	AVLKKIKHENIVLEDIYESTTHYLVMLQVSGGELFDRILRGVYTEXDASLVYQQVLS	129		
Db	72	GILKMKHPNIIQLVDVVFTRKEYFIFLELATGREVFDWILDQGYYSERDTSNVVRQVLE	131		
Qy	130	AVKYLHENGIVHRDLKPNLILYLTPEENSKIMITDFGLSKMEONGIMSTACGTPGYVAPE	189		
Db	132	AVAYLHSLKIVHRNKLLENVYVYRNKNSKIVISDFHLAKLE-NGLIKEPCGTPEYLAPE	190		
Qy	190	VLAQPKYSKAVDCWSIGVITYIILCGYPPPYEETE-----SKLFEKIKEGYEYFESP	241		
Db	191	VVGRQRYGRPVDCWAIGVIMYIILSGNPPPYEEVEEDDYENHDKNLFRLAGDYEFDSP	250		
Qy	242	FWDDISESAKDFICHLLKDPNERYTCEKALSHPWIDGNTALHROIYPSVSLQIOKNFAK	301		
Db	251	YWDDISOAAKDLVTRLMEVEQDQRTAEAEASHENISGNAASDKNIKOGVCAQIEKNFAR	310		
Qy	302	SKWQAFNAAAVVHHMKLHMLHSPGVRPEVENRPPETQASSTRPSPSEITITEAPVL	361		
Db	311	AKKKKAVRVTTL---MKRL-----RAPE--QSSTAAASASATDTATPGA	350		
Qy	362	---DHSVLPALTQLPCQHGRP-----TAPGGRSLNCLVNGSL--HSSSLVPMHQGSL	411		
Db	351	AGGATAAASGATSAPGDAARAAKSDNVAPADRSATPATDGSATPATDGSVTPATDGS	410		
Qy	412	AAGPCGCCSSCLN----IGSKGKSSYCSEPTLLKANKQNFKSEVMV---PVKASGSSH	464		
Db	411	TPATDGSVTPATDESATPATDGRATPATEESTVPTTQSSAMLATKAAATPEFAMAQPDST	470		
Qy	465	CRAQGTG	471		
Db	471	APEGATG	477		

Search completed: March 14, 2003, 14:09:39
Job time : 35 secs

[illegible]

LOCUS HS272L161 1738 bp mRNA linear PRI 21-APR-1999
 DEFINITION Human gene isolated from PAC 272L16, Chromosome 1, similar to
 ACCESSION AL049688
 VERSION AL049688.1 GI:4678721
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Rhodes.S.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-1999) E-mail contact: humquery@sanger.ac.uk
 COMMENT This sequence was generated from cDNA clones isolated using
 sequence from the bacterial clone 272L16 (AL023754) and EST data.
 The EST sequences listed match this sequence with an identity of at
 least 95% between the coordinates shown.
 Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr1/ Partial, experimentally
 determined gene.
 Sanger Centre name: dJ272L16.Cl.1.
 Location/Qualifiers
 1..1738
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1q32.1-32.3"
 /contig="1448"
 /codon_start=3
 /product="hypothetical protein"
 /protein_id="CAB41259.1"
 /db_xref="GI:4678722"
 /translation="STLEAMGRKEEDDCSSWKKOTTNIRKTFIPMEVLGSGAFSEVFL
 VKQRTGLKLFALKICKSPAFRDSLENEIAVLKKIKHENIVTLEIDYESTHYILVM
 OLVSSELFDRLRGGVYTKDASLVQVLSAVKYLHENGIVHRDLKPNLLYLP
 ENSKIMIDFDGLSKMEQNGINSTAGTPGYVAPVLAQKPYSKAVDCWSIGVITYILL
 GYPPEETESKLEPKIKGYVEFSPWDDISESAKDFICHLEKDPNERVCEKA
 LSHPWIDGNTALHRDIDYPSVLSIOKPNFASKWRQAFNAAVHHMKRLHMLHSPGV
 RPEVNRPEPQASSETSPSEFITTAPVLDHRSVALPALTPQLPCQHQRRPTAPGVR
 SLNCLVNGSLHISLSPVPMHOGSLAAGPCGCCSCLNIGSKGSSYSCSEPTLLKKANK
 KONFKSEYVMVPVKAGSGSHCRAGGTGVCILIN"
 <1..107
 /number=1
 exon
 108..236
 /number=2
 exon
 237..311
 /number=3
 exon
 312..450
 /number=4
 exon
 451..574
 /number=5
 exon
 575..650
 /number=6
 exon
 651..763
 /number=7
 exon
 764..842
 /number=8
 exon
 843..930
 /number=9
 exon
 931..1355
 /number=10
 exon
 1356..1483
 /number=11
 exon
 1484..1736
 /number=12
 exon
 432 a 491 c 458 g 357 t

ORIGIN
 Query Match 70.9%; Score 1734.8; DB 9; Length 1738;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1736; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 53 CTTCAACTCTGGAGGCAATGGGTGCGAAAGGAAGAAGATGACTGCAGTTCTTCTGGAAGAAAC 112
 Db 1 CTTCAACTCTGGAGGCAATGGGTGCGAAAGGAAGAAGATGACTGCAGTTCTTCTGGAAGAAAC 60
 QY 113 AGACCACCAACATCCGGAACACCTTCATTTTATGGAAGTGTCTGGATCAGGAGCTTTCT 172
 Db 61 AGACCACCAACATCCGGAACACCTTCATTTTATGGAAGTGTCTGGATCAGGAGCTTTCT 120
 QY 173 CAGAAGTTTTCCTGGTGAAGCAAAAGACTGACTGGGAAGCTTCTTTGCTCTCAAGTGCATCA 232
 Db 121 CAGAAGTTTTCCTGGTGAAGCAAAAGACTGACTGGGAAGCTTCTTTGCTCTCAAGTGCATCA 180
 QY 233 AGAAGTCACCTGCTTCCGGGACAGCAGCTCGAGATGAGATGCTGTGTTGAAAAAGA 292
 Db 181 AGAAGTCACCTGCTTCCGGGACAGCAGCTCGAGATGAGATGCTGTGTTGAAAAAGA 240
 QY 293 TCAAGCATGAAAACATTTGTGACCTCGAGGACATCTATGAGACACACCACCTACTACC 352
 Db 241 TCAAGCATGAAAACATTTGTGACCTCGAGGACATCTATGAGACACACCACCTACTACC 300
 QY 353 TGGTCATGACAGTTGTTTCTGGTGGGAGCTCTTTGACCGGATCTCTGGAGCGGGGTGTCT 412
 Db 301 TGGTCATGACAGTTGTTTCTGGTGGGAGCTCTTTGACCGGATCTCTGGAGCGGGGTGTCT 360
 QY 413 ACACAGAGAAGGATGCCAGCTCTGGTATCCAGCAGCTCTTGTGCGGAGTGAATACCTAC 472
 Db 361 ACACAGAGAAGGATGCCAGCTCTGGTATCCAGCAGCTCTTGTGCGGAGTGAATACCTAC 420
 QY 473 ATGAGATGCGATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTG 532
 Db 421 ATGAGATGCGATCGTCCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTG 480
 QY 533 AAGAACTCTAAGATCATGATCAGTCTTGTGCTGTCCAAAGATGGAACAGAAATGGCA 592
 Db 481 AAGAACTCTAAGATCATGATCAGTCTTGTGCTGTCCAAAGATGGAACAGAAATGGCA 540
 QY 593 TCATGTCACCTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCCGCAAAAC 652
 Db 541 TCATGTCACCTGCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTCTGGCCCGCAAAAC 600
 QY 653 CCTACAGAAAGGCTGTGGATTTGCTGCTCCATCGGCGTCATCCTACATATCTGTG 712
 Db 601 CCTACAGAAAGGCTGTGGATTTGCTGCTCCATCGGCGTCATCCTACATATCTGTG 660
 QY 713 GATACCCCGCTTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCT 772
 Db 661 GATACCCCGCTTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCT 720
 QY 773 ACTATGAGTTTGTAGTCTCCATTTCTGGGATGACATTTCTAGTCAGCAAGGACTTTATTT 832
 Db 721 ACTATGAGTTTGTAGTCTCCATTTCTGGGATGACATTTCTAGTCAGCAAGGACTTTATTT 780
 QY 833 GCCACTTGTGTAGAGAGATCCGAAACAGCGGTACCTGTGAGAGGCGCTTGAAGTCATC 892
 Db 781 GCCACTTGTGTAGAGAGATCCGAAACAGCGGTACCTGTGAGAGGCGCTTGAAGTCATC 840
 QY 893 CCTGATTTAGCGAAACACAGCGCTCCACCGGACATCTACCCATCAGTCAGCCTCCAGA 952
 Db 841 CCTGATTTAGCGAAACACAGCGCTCCACCGGACATCTACCCATCAGTCAGCCTCCAGA 900
 QY 953 TCCAGAAAGAACTTTGTCTAAGCAAGTGGAGCAAGCTTCAACGCGAGCAGCTGTGGTGC 1012
 Db 901 TCCAGAAAGAACTTTGTCTAAGCAAGTGGAGCAAGCTTCAACGCGAGCAGCTGTGGTGC 960
 QY 1013 ACCACATGAGAGAGCTACACATGAACCTGCACAGCGCGGCGCTCCGCCAGAGGTGAGA 1072
 Db 961 ACCACATGAGAGAGCTACACATGAACCTGCACAGCGCGGCGCTCCGCCAGAGGTGAGA 1020

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 41 Row: i Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction.

Location/Qualifiers
1. 2416

```
/sub_node= D110D
/note="Vector: pCMV-SPORT6"
```

Query Match 54.3%; Score 1328.6; DB 10; Length 2416;
Best Local Similarity 76.1%; Pred. No. 0;
Matches 1853; Conservative 0; Mismatches 494; Indels 88; Gaps 14;

QY 91 GACTGCAGTTCCTGGGAAGAAACAGACCACCAACATCCGGAAAACCTTCATTTTATGGAA 150

QY 151 GTGCTGGGATCAGGAGCTTCTCAGAAAGTTTCCCTGGTGAAGCAAAGACTGACTGGGAAG 210

OV 211 CTCTTTGCTCTGAAGTGCATCAAGAAAGTCACCTGGCCTTCCGGGACAGCAGCCTGGAGAAT 270

QV 271 CACATTCCTGTTCAAAACAATCAGCATATCATCGCCGCATCATF 330

331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

C	
A	E
E	C
C	C
C	C
C	C
E	C
C	E
C	E
C	E
C	E
C	C
E	C
C	C
C	C
C	C
E	C
E	C
C	C
C	C
C	C
C	C
E	C
E	C
C	C
C	C
C	C
307	

[illegible]

BC021840

DEFINITION

KEYWORDS

ORGANIZATIONAL

REFERENCES

TITLE

REMARK

HS272L161 1738 bp mRNA linear PRI 21-APR-1999
LOCUS Human gene isolated from PAC 272L16, chromosome 1, similar to
DEFINITION calcium/calmodulin dependent protein kinases.
ACCESSION AL049688
VERSION AL049688.1 GI:4678721
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia; Eutheria: Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1738)
Rhodes, S.
Direct Submission
TITLE Submitted (21-APR-1999) E-mail contact: humquery@sanger.ac.uk
JOURNAL This sequence was generated from cDNA clones isolated using
COMMENT sequence from the bacterial clone 272L16 (AL023754) and EST data.
The EST sequences listed match this sequence with an identity of at
least 95% between the coordinates shown.
Further information can be found at
http://www.sanger.ac.uk/Hgp/Chri/ Partial, experimentally
determined gene.
Sanger Centre name: G0272L16.C1.1.
Location/Qualifiers
1..1738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1q32.1-32.3"
/codon_start=3
/product="hypoetical protein"
/protein_id="CAB41259.1"
/db_xref="GI:4678722"
/translation="STLEAMGRKEEDDCSSWKQNTNIRKTFIPWVLGSGAFSEVPL
VKQLTKLPALCKIKSPAFRDSSLENIATLVKIKKHENIVTLEDIYESPTHYLYVM
QLVSGGELEDRILIERGYTEKDALVIQOVSARKYLHENGIVHRDKPENLYLYTPE
ENSKMITDGLSKMEQNGIMSTAGTGYVAPEYLAQKPYSKAVDCWSIGVITYILL
GYPPFEYETKLFKIKEGYEFESPPWDDISESKDFICHLEKDPNERYTCERA
LSHPIDNTALHRDIPYPSVLIQIKNFPAKSKWROAFNAAVHVHMKLHMNLSPGV
RPEVNPPTQASSETSRPSPSEITITTEAPVLDHVALPALTPCOHGRPTAPGGR
SLNGLSHLSHSSLYPMHQSGLAAGFCGCGSSCLNTGSKGSKSSSEPTLLKKANK
KQNFKSYNVPVVRASGSSCRAGGTGCLIM"
/number=1
108..236
/number=2
237..311
/number=3
312..450
/number=4
451..574
/number=5
575..650
/number=6
651..763
/number=7
764..842
/number=8
843..930
/number=9
931..1355
/number=10
1356..1483
/number=11
1484..1736
/number=12
misc_feature
/note="matches EST A1215131 from clone IMAGE:1925595"
exon
exon
exon
exon
exon
exon
exon
misc_feature
exon
exon
exon
misc_feature
exon
exon
misc_feature
exon
exon
misc_feature
exon

BASE COUNT 432 a 491 c 458 g 357 t
ORIGIN
Alignment Scores:
Pred. No.: 2,33e-170 Length: 1738
Score: 2513.00 Matches: 476
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-960-643-2 (1-476) x HS272L161 (1-1738)
QY 1 MetGlyArgLysGluGluAspAspCysSerTrpLysLysGlnThrThrAsnLeuArg 20
Db 18 ATGGGTGCGAAGGAAGAGATGACTGCGAGTTCCTGGAAGAACAGACCAACATCCGG 77
QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
Db 78 AAACCTTCCTTTTATGGAAGTCTGGATCAGAGCTTCTCGAAGTTTCTCGGGTG 137
QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysLysLysLysSerProAlaPhe 60
Db 138 AACCAAGACTGACTGGGAAGCTCTTGTCTCTGAAGTCATCAAGAAGTCACTGCCTTC 197
QY 61 ArgAspSerSerLeuGluAsnGluLeuAlaValLeuLysLysLysHisGluAsnLeu 80
Db 198 CGGACAGCAGCTGGAGATGAGATGCTGTGTTGAAAAGATCAAGCATGAAACATT 257
QY 81 ValThrLeuGluAspIleTrpGluSerThrThrHisTrpLysLysLysLysValMetGlnLeuVal 100
Db 258 GTGACCTGGAGACATCTATGAGAGCACCACCCACTACTACCTGGTGCATGCAGCTTGT 317
QY 101 SerGlyGlyLeuPheAspArgIleLeuGluArgGlyValThrGlnLysAsnAla 120
Db 318 TCTGGTGGGAGCTCTTTCACCGGATCTGGAGCGGGGTCTACACAGAGAAGATGCC 377
QY 121 SerLeuValIleGlnValLeuSerAlaValLysLysLysLysGluAsnGlyVal 140
Db 378 AGTCTGGTATCCAGCAGCTCTTGTGCGAGTGAATACCTACATGAGATGGCATCGTC 437
QY 141 HisArgAspLeuLysProGluAsnLeuLeuThrProGluGluAsnSerLysIle 160
Db 438 CACAGAGACTTAAAGCCGAAACCTGCTTACCTTACCTCCCTGAAGAGAACTCTAAGATC 497
QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
Db 498 ATCATCATGACTTGGTCTCTCCAAAGATGGAACAGAGATGGCATCATGTCACCTGCCTGT 557
QY 181 GlyThrProGlyTrpValAlaProGluValLeuAlaGlnLysProTrpSerLysAlaVal 200
Db 558 GGGACCCAGGCTACGTGGCTCCAGAGTCTGGCCAGAACCCCTACAGACAGGCTGTG 617
QY 201 AspCysTrpSerIleGlyValIleThrTrpIleLeuLysGlyTrpProPheTrp 220
Db 618 GATTGCTGGTCCATCGCGCTCATCACCACATATTTGCTCTGTGGATACCCCATCTAT 677
QY 221 GluGluThrGluSerLysLeuPheGluLysLysGluGlyTrpGluPheGluSer 240
Db 678 GAAGAAACGAGGCTTAAGCTTTTCGAGAGAGATCAAGAGAGGCTTACTATGAGTTTGAGTCT 737
QY 241 PropheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
Db 738 CCATTCGGGATGACATTTCTGAGTCAGCAAGGACTTATTTGCCACTTCTGTTGAGAG 797
QY 261 AspProAsnGluArgTrpThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
Db 798 GATCCGAACGAGCGGTACACCTGTGAGAAGGCTTGTGAGTCATCCCTGGATTGACGGAAC 857
QY 281 ThrAlaLeuHisArgAspIleTrpProSerValSerLeuGlnIleGlnLysAsnPheAla 300
Db 858 ACAGCCCTCCACCGGACATCTACCATCAGTCAGCTCCAGCTCCAGAGAACTTTGCT 917

QY 301 LysSerLysTyrAlaGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320
 |||||
 Db 918 AGAGCAAGTGGAGGAGGCTTCACGACGAGCTGGTGGTCCACACATGAGGAAGCTA 977
 |||||
 QY 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProGluThr 340
 |||||
 Db 978 CACATGAACCTGCACACGCCGCGCTCCGCCAGAGGTGGAGAACAGCGCCCTGAAACT 1037
 |||||
 QY 341 GlnAlaSerGluThrSerArgProSerProGluThrLeuThrLeuThrGluAlaProVal 360
 |||||
 Db 1038 CAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCAGGAGGACCTGTC 1097
 |||||
 QY 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
 |||||
 Db 1098 CTGGACCAACAGTGTAGACATCCCTGCCCTGACCCCAATACCTGCCAGCATGCCGCCCG 1157
 |||||
 QY 381 ProThrAlaProGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
 |||||
 Db 1158 CCACCTGCCCTGGTGGCAGTCCCTCAACTGCTGGTCAATGCTCCCTCCACATCAGC 1217
 |||||
 QY 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420
 |||||
 Db 1218 AGCAGCTGGTGGCCATGATCAGGGTCCCTGGCGCGCGGCCCTGTGCTGCTGCTC 1277
 |||||
 QY 421 SerCysLeuAsnIleGlySerLysGlyLysSerSerTyrCysSerGluProThrLeuLeu 440
 |||||
 Db 1278 AGCTGCCCTGAACATGGGACAAAGAAAGTCTCTCTACTGCTGAGCCACACATCCTC 1337
 |||||
 QY 441 LysLysAlaAsnLysLysGlnAsnPheLysSerGluValMetValProValLysAlaSer 460
 |||||
 Db 1338 AAAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTGATGGTACCAAGTAAAGCCAGT 1397
 |||||
 QY 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
 |||||
 Db 1398 GGCAGCTCCCACTGCCGGGAGGACAGTGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1445
 |||||

RESULT 2
 AX399682
 LOCUS AX399682 2447 bp DNA linear PAT 06-JUN-2002
 DEFINITION Sequence 3 from Patent WO0224947.
 ACCESSION AX399682
 VERSION AX399682.1 GI:21335455
 KEYWORDS human.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Delaney, A.D. and Yoganathan, T.
 Cancer associated protein kinases and their uses
 Patent: WO 0224947-A 3 28-MAR-2002;
 KINETEK PHARMACEUTICALS INC (CA); UNIV BRITISH COLUMBIA (CA)
 Location/Qualifiers
 1..2447
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 70..1500
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD33521.1"
 /db_xref="GI:21335456"
 /translation="MGRKEEDSCSWKKOTTNIRKTFIPMEVLGSGAFSEVLYKQRL
 TGLFKALKIKKSPAFROSSLLENIAVLKKIKHENIVTLDIEDYESTHYLYVMQVLSG
 GELFDRLIRGVYTERDASLVQVLSAVLYKHENGIVHRDLKPENLVLYTPEENSKI
 MITDFGLSKMEQNGIMSTACGTPGVVAQKPYKSAWDCWSIGVITYILLGCGPP
 FYEETESKLPEIKKEGYIEFFSPWDDIISAKDFICHLEKDPNRYCTCEKALSHPW
 IDGNTALRDYIPSVSLQIKNFPAKSKWQAFNAVAVVMHMKLHMNLSHPCVPEVE
 NRPEQASSETSPSSPETTEAPVLDHVALPALTQLPCQGRPPPTAPGRSLNCL
 VNSLHSSSLPMMQGLAGPCCGCCSCLNIGSRKSTCYSEPTLLKANKKNKFNK
 SEVVPVPAKSGSSHCHRGAGTGVGLIM"

FEATURES
 source
 CDS
 70..1500
 /note="unnamed protein product"
 /protein_id="CAD33521.1"
 /db_xref="GI:21335456"
 /translation="MGRKEEDSCSWKKOTTNIRKTFIPMEVLGSGAFSEVLYKQRL
 TGLFKALKIKKSPAFROSSLLENIAVLKKIKHENIVTLDIEDYESTHYLYVMQVLSG
 GELFDRLIRGVYTERDASLVQVLSAVLYKHENGIVHRDLKPENLVLYTPEENSKI
 MITDFGLSKMEQNGIMSTACGTPGVVAQKPYKSAWDCWSIGVITYILLGCGPP
 FYEETESKLPEIKKEGYIEFFSPWDDIISAKDFICHLEKDPNRYCTCEKALSHPW
 IDGNTALRDYIPSVSLQIKNFPAKSKWQAFNAVAVVMHMKLHMNLSHPCVPEVE
 NRPEQASSETSPSSPETTEAPVLDHVALPALTQLPCQGRPPPTAPGRSLNCL
 VNSLHSSSLPMMQGLAGPCCGCCSCLNIGSRKSTCYSEPTLLKANKKNKFNK
 SEVVPVPAKSGSSHCHRGAGTGVGLIM"

BASE COUNT 590 a 707 c 604 g 546 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,52e-170 Length: 2447
 Score: 2513.00 Matches: 476
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-960-643-2 (1-476) x AX399682 (1-2447)

QY 1 MetGlyArgLysGluLysGluLysCysSerSerTrpLysGluThrThrAsnIleArg 20
 |||||
 Db 70 ATGGTCCAAAGAGGAGAGATGATGCTGAGTCTCTGGAGAAACACACCAACATCCGG 129
 |||||
 QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
 |||||
 Db 130 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAAGTTTCTCGTG 189
 |||||
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
 |||||
 Db 190 AACCAAAAGACTGCTGGAGAGCTTTTCCCTGAACTGCTCAAGAGTCAACAGTCCCTC 249
 |||||
 QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysIleLysHisGluAsnIle 80
 |||||
 Db 250 CGGGACAGCAGCCCTGGAGATGATGCTGCTTGGAAAAAGATCAAGCATGAACATT 309
 |||||
 QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGluLeuVal 100
 |||||
 Db 310 GTGACCTTGGAGGACATCTATGAGAGCACCACCACTACTACCTGCTGTCATGCGCT 369
 |||||
 QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
 |||||
 Db 370 TCTGGTGGGAGACTCTTTGACCGATCTTGAGCGGGGTCTACACAGAGAGAGATGCC 429
 |||||
 QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
 |||||
 Db 430 AGCTCGGTGATCCAGCAGGCTTGTGCGCAGTGAATACTACATGAGAATGGCATCGTC 489
 |||||
 QY 141 HisArgAspLeuLysProGluAsnLeuTyrLeuThrProGluGluAsnSerLysIle 160
 |||||
 Db 490 CACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCCCTGAAGAACTCTAAGATC 549
 |||||
 QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
 |||||
 Db 550 ATGATCATGCTATTGGTCTGCTCCAGATGGAACACAAATGGCATCATGCTCCACCT 609
 |||||
 QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
 |||||
 Db 610 GGGACCCAGGCTAGTGGCTCCAGAAAGTCTGGCCAGAAACCTTACAGCAAGGCTGTG 669
 |||||
 QY 201 AspCysTyrSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProPheTyr 220
 |||||
 Db 670 GATTGCTGCTCATCGGCGCTCATCACCTACATATTGCTCTGTGATATACCCCGCTCTAT 729
 |||||
 QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
 |||||
 Db 730 GAAGAAACGGAGCTCTAGCTTTTCGAGAAGATCAGAGAGGCTACTATGATGTTGAGTCT 789
 |||||
 QY 241 ProPheTrpAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
 |||||
 Db 790 CCATTTCTGGATGATCATTTCTGAGTACGCAAGGACTTTATTTGGCACTTCTGTTGAAAG 849
 |||||
 QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
 |||||
 Db 850 GATCCGAAGAGCGGTACACCTGTGGAAGGCCCTTGGATCATCCCTGGATGACGGAAC 905
 |||||
 QY 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
 |||||
 Db 910 AGGCCCCCTCCACCGGAGATCTACCCATCAGTCCAGCTCCAGATCCAGAAAGATTTGCT 969
 |||||
 QY 301 LysSerLysTyrArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320
 |||||
 Db 970 AAGAGCAAGTGGAGGCAAGCCTTCAACGCGAGCAGCTGTGTGTCACCATGAGGAAGCTA 1029
 |||||


```
Db 1621 AGCAGGAGCAGTTCTGGCCAGAGCACCAGCTGTCGCCAGGGGAGCCCTCATAG 1680
QY 1681 GAGGCCAGAGGAGGCCAGCCAGGCTAGAGCCTTTGTAAGCTGTGAGCAGGAGAGC 1740
Db 1681 GAGGCCAGAGGAGGCCAGCCAGGCTAGAGCCTTTGTAAGCTGTGAGCAGGAGAGC 1740
QY 1741 GGTGCCACACAGCTTCCAGCTGCCCTGACCTGACCTGCTATGCCGCCACACCTACGTG 1800
Db 1741 GGTGCCACACAGCTTCCAGCTGCCCTGACCTGACCTGCTATGCCGCCACACCTACGTG 1800
QY 1801 CCGTGGCTCTGTCAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 1860
Db 1801 CCGTGGCTCTGTCAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 1860
QY 1861 AAAGCTTAATGGGCTGGCCAGGCTGTGTCACCTTCTCCAGCAAGCAAGCCATATGGAGCATC 1920
Db 1861 AAAGCTTAATGGGCTGGCCAGGCTGTGTCACCTTCTCCAGCAAGCAAGCCATATGGAGCATC 1920
QY 1921 TACCCAGACTCCCACTCTGCACACACTCACTCCCACTCTCAAGCTCTCAAGCTCTTGGC 1980
Db 1921 TACCCAGACTCCCACTCTGCACACACTCACTCCCACTCTCAAGCTCTCAAGCTCTTGGC 1980
QY 1981 CAGATTGGGCTCAATTAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 1981 CAGATTGGGCTCAATTAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
QY 2041 TGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 2041 TGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
QY 2101 TGGCGTGAGACCCAGCAGCAGCAGCTTCTGGCGCTTCTGCTGCTGCTGCTGCTGCTGCT 2160
Db 2101 TGGCGTGAGACCCAGCAGCAGCAGCTTCTGGCGCTTCTGCTGCTGCTGCTGCTGCTGCT 2160
QY 2161 GCCACACCTTCCAAAGTGGATAGAAAGTTCATGATGATGATGATGATGATGATGATGAT 2220
Db 2161 GCCACACCTTCCAAAGTGGATAGAAAGTTCATGATGATGATGATGATGATGATGATGAT 2220
QY 2221 ATCTGCGCCACATGCTCTCGTGCACACACACCAATGAGTGAAGTGAAGTGAAGTGAAGT 2280
Db 2221 ATCTGCGCCACATGCTCTCGTGCACACACACCAATGAGTGAAGTGAAGTGAAGTGAAGT 2280
QY 2281 TTTAATGCTGCGCAGAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db 2281 TTTAATGCTGCGCAGAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
QY 2341 AGCACACATTCCTGCTTCCCGAGTTCCTGCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCT 2400
Db 2341 AGCACACATTCCTGCTTCCCGAGTTCCTGCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCT 2400
QY 2401 AGTGTGCTGCTTAATAAATGGACATATTTTCTCTAATAAATAAATAA 2447
Db 2401 AGTGTGCTGCTTAATAAATGGACATATTTTCTCTAATAAATAAATAA 2447
```

RESULT 2

AAI60703

ID AAI60703 standard; cDNA; 1956 BP.

XX AC AAI60703;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 4692.

XX XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

XX KW peripheral nervous system; neuropathy; central nervous system; CNS;

XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

XX KW leukaemia; ss.

OS Homo sapiens.

XX WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX XX WPI; 2001-442253/47.

XX DR P-PSDB; AAM41547.

XX XX Novel nucleic acids and polypeptides, useful for treating disorders

XX PT such as central nervous system injuries -

XX PS Claim 1; SEQ ID NO 4692; 10078pp; English.

XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful

XX CC in gene therapy. A composition containing a polypeptide or polynucleotide

XX CC of the invention may be used to treat diseases of the peripheral nervous

XX CC system, such as peripheral nervous injuries, peripheral neuropathy and

XX CC localised neuropathies and central nervous system diseases, such as

XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX CC activation of the activities such as: immune system suppression,

XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

XX CC assays for receptor activity, arthritis and inflammation, leukaemias and

XX CC C.N.S disorders.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification.

XX SQ Sequence 1956 BP; 474 A; 559 C; 508 G; 414 T; 1 other;

Query Match

Best Local Similarity 78.3%; Score 1915; DB 22; Length 1956;

Matches 1941; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY 12 CTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTCTGAGAGCTTCAACTCTGGAGGCAAT 71

Db 7 CTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTCTGAGAGCTTCAACTCTGGAGGCAAT 66

QY 72 GGGTCGAAGGAAGAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 131

Db 67 GGGTCGAAGGAAGAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 126

QY 132 AACCTTCATTTTATGGAAGTCTGGGATGATGATGATGATGATGATGATGATGATGATGAT 191

Db 127 AACCTTCATTTTATGGAAGTCTGGGATGATGATGATGATGATGATGATGATGATGATGAT 186

QY 192 GCAAGAGTACTGCTGGGAGTCTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 251

Db 187 GCAAGAGTACTGCTGGGAGTCTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 246

QY 252 GGACAGCAGCCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 311

Db 247 GGACAGCAGCCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 306

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 17:36:01 ; Search time 328 Seconds
(without alignments)
3268.144 Million cell updates/sec

Title: US-09-960-643-2

Perfect score: 2513

Sequence: 1 MGRKEDDCSSWKQTTNIR.....VKAGSSHCRAGQTGVCLIM 476

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Xgapop 10.0 , Ygapext 0.5	
Ygapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O/cgn2_1/USPTO.spool/US09960643/runat_07032003_090517_19867/app.query.fasta_1.647
-DB=N_Geneseq_101002 -QPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09960643.qcgn_1_1_263_arunat_07032003_090517_19867 -NCPU=6 -ICPU=3
-NO_XLPAY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -FPGAPEXT=7
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FPGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DLOOP=6 -DELEXT=7

Database :

1: N_Geneseq_101002.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
25: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2513	100.0	1956	22	Human polynucleoti
2	2513	100.0	2447	24	Human calmodulin k
3	2499.5	99.5	2165	22	Human polynucleoti
4	2327.5	92.6	2689	22	Human diagnostic a
5	1250.5	49.8	1736	22	Human kinase (PKIN
6	1246	49.6	1074	22	Polynucleotide seq
7	1246	49.6	1074	22	Nucleotide sequenc
8	1246	49.6	1158	22	Nucleotide sequenc
9	1246	49.6	1565	24	Human Cam kinase I
10	1246	49.6	1661	22	Human kinase (PKIN
11	1246	49.6	1671	22	Nucleotide sequenc
12	1246	49.6	2164	22	Human polynucleoti
13	1243	49.5	1733	22	Human polynucleoti
14	1238.5	49.3	1694	22	Human polynucleoti
15	1237	49.2	1578	22	Human death domain
16	1150.5	45.8	1728	21	cdNA sequence enco
17	1070	42.6	1458	19	Calmodulin-depende
18	1041	41.4	1032	22	Polynucleotide seq
19	1036	41.2	1512	23	DNA encoding novel
20	990	39.4	1282	20	Disease associated
21	957.5	38.1	2264	23	Drosophila melanog
22	753	30.0	2494	22	Human full-length
23	753	30.0	2747	22	Human polynucleoti
24	753	30.0	2897	22	Human protein phos
25	753	30.0	2905	22	Human full-length
26	747.5	29.7	2297	22	Nucleotide sequenc
27	747.5	29.7	3017	22	Human kinase (PKIN
28	747.5	29.7	3124	24	Human protein kin
29	747	29.7	2840	22	Human polynucleoti
30	733.5	29.2	3022	22	Human full-length
31	726.5	28.9	1572	22	Human protein enco
32	683.5	27.2	1774	23	DNA encoding novel
33	683.5	27.2	1774	23	DNA encoding novel
34	667	26.5	1763	23	DNA encoding novel
35	664.5	26.4	3471	18	PSKH-1 cDNA. Homo
36	661.5	26.3	1764	23	DNA encoding novel
37	650.5	25.9	1386	22	Human polynucleoti
38	650.5	25.9	1450	22	Human polynucleoti
39	639.5	25.4	3552	24	Human serine/threo
40	633.5	25.2	2667	24	Human PKIN-14 cDNA
41	625.5	24.9	2301	24	Human serine/threo
42	615.5	24.5	2110	21	cdNA sequence enco
43	611.5	24.3	1803	22	cdNA encoding huma
44	611.5	24.3	1803	24	Human kinase (PKIN
45	611.5	24.3	7626	24	Human ovary specif

ALIGNMENTS

RESULT 1

AAI60703

ID AAI60703 standard; cdNA; 1956 BP.

XX AAI60703;

AC AAI60703;

DT 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 4692.

XX Human; notoprotic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.

XX W0200153312-A1.
 XX 26-JUL-2001.
 XX 26-DEC-2000; 2000WQ-US34263.
 XX 21-JAN-2000; 2000US-0488725.
 XX 25-APR-2000; 2000US-0552317.
 XX 09-JUL-2000; 2000US-0598042.
 XX 19-JUL-2000; 2000US-0620312.
 XX 03-AUG-2000; 2000US-0653450.
 XX 14-SEP-2000; 2000US-0662191.
 XX 19-OCT-2000; 2000US-0693036.
 XX 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPT; 2001-442253/47.
 XX P-PSDB; AAM41547.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 XX such as central nervous system injuries -
 XX Claim 1; SEQ ID NO 4692; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 XX immunosuppressant and cytostatic activity. The polynucleotides are useful
 XX in gene therapy. A composition containing a polypeptide or polynucleotide
 XX of the invention may be used to treat diseases of the peripheral nervous
 XX system, such as peripheral nervous injuries, peripheral neuropathy and
 XX localised neuropathies and central nervous system diseases, such as
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 XX lateral sclerosis, and Shy-brager Syndrome. Other uses include the
 XX utilisation of the activities such as: Immune system suppression,
 XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 XX assays for receptor activity, cancer diagnosis and inflammation, leukaemias and
 XX C.N.S disorders.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification.
 XX SQ Sequence 1956 BP; 474 A; 559 C; 508 G; 414 T; 1 other;
 Alignment Scores:
 Pred. No.: 5.24e-206 Length: 1956
 Score: 2513.00 Matches: 476
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-960-643-2 (1-476) x AAI60703 (1-1956)
 QY 1 MetGlyArgLysGluGluAspAspCysSerSerTrpLysLysGlnThrThrAsnLeuArg 20
 DB 65 ATGGGTGGAAGAGGAGAGATGATGCTGAGTTCTTGAAGAAACAGACCAACATCCGG 124
 QY 21 LysThrPheLeuPheMetGluValLeuGlySerGlyValAlaPheSerGluValPheLeuVal 40
 DB 125 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGACCTTTCAGAACTTTCTCGTG 184
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysLysLysSerProAlaPhe 60
 DB 185 AAGCAAGAACTGACTGGGAAGCTTTTGTCTCTGAAGTGCATCAAGAGTCACTGCTTC 244
 QY 61 ArgAspSerSerLeuGluAsnGlnLeuAlaValLeuLysLysLysLysGlnAsnLeu 80
 DB 1325 AGCTGCTGAACATTTGGGAGCAAGGAAAGTCTCTCTACTGCTCTGAGACCACTCTCTC 1384

DB 245 CGGGACAGCAGCCTGGAGATGAGATTGCTGCTGTAAGAAAGATCAAGCATGAAACATT 304
 QY 81 ValThrLeuGluAspPheLeuSerThrThrHisThrLysLeuValMetGlnLeuVal 100
 DB 305 GTGACCCCTGGAGGACATCTATGAGAGCACCACCCTACTTACCTGGTCAATGACCTTGT 364
 QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValThrGluLysAspAla 120
 DB 365 TCTGTGGGGAGCTCTTTGACCGGATCTCTGAGCGGGGTCTCTACACAGAGAGGATGCC 424
 QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysThrLeuHisGluAsnGlyIleVal 140
 DB 425 AGTCGGGTGATCCAGCAGGCTTGTGCGAGTGAATACCTACTAGAAATGGCATCGTC 484
 QY 141 HisArgAspLeuLysProGluAsnLeuLeuThrProGluGluAsnSerLysIle 160
 DB 485 CACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCCTGAAGAGAACTTAAGATC 544
 QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
 DB 545 ATGATCACTGACTTGTGCTCTCAAGATGGACAGAAATGGCATCTGCTCCACTGCTGT 604
 QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
 DB 605 GGGACCCAGGGCTAGCTGGCTCCAGAAAGTGTGCGCCAGAAACCCCTACAGCAAGGCTGTG 664
 QY 201 AspCysTTPSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220
 DB 665 GATTGCTGCTCAGCGGCTCATCCTACATATTGCTGTGGATACCCCTCATTTCTAT 724
 QY 221 GluGluThrGluSerLysLeuPheGluLysLysLysGlyTyrTyrGluPheGluSer 240
 DB 725 GAAGAAACGGAGCTCTTAAGCTTTTCGAGAAGATCAAGAGGGGCTACTATGATGTTGAGTCT 784
 QY 241 ProPheTTPAspPheLeuSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
 DB 785 CCATCTCTGGGATGACATCTTCAGTACAGCAAGGACCTTATTTGCCACTTGTCTGGAGAAG 844
 QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProTrpIleAspGlyAsn 280
 DB 845 GATCCGAAAGAGCGGTACACCTGTGAGAAGGCCCTTGAGTATCCCTGGATTCAGCGAAAC 904
 QY 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
 DB 905 ACAGCCCTCCACCGGACATCTACCCATCAGCTCCAGCTCCAGATCCAGAGAAGACTTGTCT 964
 QY 301 LysSerLysTTPArgGlnAlaPheAsnAlaAlaValHisHisMetArgLysLeu 320
 DB 965 AAGAGCAAGTGGAGCAAGCCCTTCAACGACGACGCTGTGTGACCAACATCAGCAAGCTA 1024
 QY 321 HisMetAsnLeuHisSerProGlyValArgProGluValGluAsnArgProGluThr 340
 DB 1025 CACATGAACCTGCACAGCCCGGGGCTCCGCCAGAGTGGAGAACAGCCGCTGAAACT 1084
 QY 341 GlnAlaSerGluThrSerArgProSerProGluLleThrIleThrGluAlaProVal 360
 DB 1085 CAAGCCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCACCATCACCAGGACCTGTCT 1144
 QY 361 LeuAspHisSerValAlaLeuProAlaLeuThrGlnLeuProCysGlnHisGlyArgArg 380
 DB 1145 CTGGACACAGGTGTAGCACTCCCTGCGCTGACCCCAATTTACCTGCCAGCATGGCCGCGG 1204
 QY 381 ProThrAlaProGlyGlyArgSerLeuAsnCysLeuValAsnGlySerLeuHisIleSer 400
 DB 1205 CCCACTGCCCTGTGTGGCAGGTCCCTCAACTGCCTGCTCAATGCTCCCTCCACATCAGC 1264
 QY 401 SerSerLeuValProMetHisGlnGlySerLeuAlaAlaGlyProCysGlyCysSer 420
 DB 1265 AGCAGCCTGTGCTCCCATGCATCAGGGGTCCCTGGCCGCCGCGGCTGTGCTGTCTTC 1324
 QY 421 SerCysLeuAsnIleGlySerLysLysSerSerTyrCysSerGluProThrLeuLeu 440
 DB 1325 AGCTGCTGAACATTTGGGAGCAAGGAAAGTCTCTCTACTGCTCTGAGACCACTCTCTC 1384

QY 441 LysLysAlaAsnLysLysGlnAsnPhelYsSerGluValMetValProValLysAlaSer 460
 Db 1385 AAAAGGCCACAAAAACAGAACTTCAAGTCGGAGGTCAATGGTACCAGTTAAAGCCAGT 1444
 QY 461 GlySerSerHisCysArgAlaGlyGlnThrGlyValCysLeuIleMet 476
 Db 1445 GGCAGCTCCCACTGCCGGCAGGCAGACTGGAGTCTGTCTCATTTNG 1492

RESULT 2

AD36140
 ID AAD36140 standard; DNA; 2447 BP.

XX AC AAD36140;

XX DT 09-AUG-2002 (first entry)

XX DE Human calmodulin kinase, CAMK-X1 gene.

XX KW Human; cytostatic; antisenase gene therapy; screening; protein kinase;
 KW cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-X1;
 KW calmodulin kinase; enzyme; gene; chromosome lq32.1-32.3; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 70..1500

FT FT /*tag= a

FT FT /product= "Human CAMK-X1 protein"

XX PN WO200224947-A2.

XX PD 28-MAR-2002.

XX PF 20-SEP-2001; 2001WO-IB02237.

XX PR 20-SEP-2000; 2000US-233999P.

XX PR 02-OCT-2000; 2000US-237419P.

XX PR 02-OCT-2000; 2000US-237423P.

XX PR 04-OCT-2000; 2000US-238558P.

XX PR 10-MAY-2001; 2001US-290555P.

XX (KINE-) KINETEK PHARM INC.

PA (UYBR-) UNIV BRITISH COLUMBIA.

XX Yoganathan T, Delaney AD;

XX WPI; 2002-394145/42.

DR P-PSDB; AAE22764.

PT Diagnosing cancer, comprises determining the upregulation of expression
 of a nucleic acid sequence encoding a protein kinase or upregulation of
 expression of the protein kinase, in the cancer

PS Claim 16; Page 62-64; 87pp; English.

XX The invention relates to a method for screening biologically active agent
 CC that modulates cancer associated protein kinase function. The invention
 CC also relates to a method for diagnosing cancer comprising determining the
 CC upregulation of expression of a nucleic acid sequence encoding a protein
 CC kinase. The method is useful for diagnosing cancer. A protein kinase is
 CC useful for screening biological agents that modulate cancer associated
 CC protein kinase function. Downregulating the activity of protein kinase is
 CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon
 CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy
 CC derived tumours and inflammatory samples such as arthritic synovium, for
 CC amplified DNA in the cell or increased expression of corresponding mRNA
 CC or protein and is also useful to detect differences in expression levels
 CC such as molecular weight, amino acid and nucleotide sequences between the
 CC two cells. The present sequence is human calmodulin kinase CAMK-X1 gene
 CC located on chromosome lq32.1-32.3.

XX SQ Sequence 2447 BP; 590 A; 707 C; 604 G; 546 T; 0 other;

Alignment Scores:
 Pred. No.: 7.2e-206 Length: 2447
 Score: 2513.00 Matches: 476
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-09-960-643-2 (1-476) x AAD36140 (1-2447)

QY 1 MetGlyArgLysGluAspAspCysSerSerTyrLysLysGlnThrThrAsnIleArg 20
 Db 70 ATGGGTGAAAGGAAGAGATGACTGCAGTTCTCTGGAAGAAACAGACCACCAACATCCGG 129
 QY 21 LysThrPheIlePheMetGluValLeuGlySerGlyAlaPheSerGluValPheLeuVal 40
 Db 130 AAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAGATTTTCCCTGGTG 189
 QY 41 LysGlnArgLeuThrGlyLysLeuPheAlaLeuLysCysIleLysLysSerProAlaPhe 60
 Db 190 AAGCAACACACTGACTGGGAAGCTCTTGTGCTGAAGTGCATCAAGAGTCACTGCCTTC 249
 QY 61 ArgAspSerSerLeuGluAsnGluIleAlaValLeuLysLysIleLysHisGluAsnIle 80
 Db 250 CGGGACACAGCCCTGGAGATGAGATTGCTGTGTTGAAAGATCAAGACATGAAACAT 309
 QY 81 ValThrLeuGluAspIleTyrGluSerThrThrHisTyrTyrLeuValMetGlnLeuVal 100
 Db 310 GTGACCTTGGAGGACATCTATGAGAGCACACACCTACTACTGCTGTCATGAGCTGTT 369
 QY 101 SerGlyGlyGluLeuPheAspArgIleLeuGluArgGlyValTyrThrGluLysAspAla 120
 Db 370 TCTGTGGGGAGCTCTTTGACCGGATCTGTGAGCGGGGTGTACACAGAGAAAGATGCC 429
 QY 121 SerLeuValIleGlnGlnValLeuSerAlaValLysTyrLeuHisGluAsnGlyIleVal 140
 Db 430 AGTCGTGGTGTCCAGCAGGTCTTGTGCGCAGTGAATACCTCATGAGAAATGGCATCGTC 489
 QY 141 HisArgAspLeuLysProGluAsnLeuLeuTyrLeuThrProGluGluAsnSerLysIle 160
 Db 490 CACAGAGACTTAAAGCCGCGAAAGCTGCTTTACCTTACCTTACCTTACCTTACCTTAC 549
 QY 161 MetIleThrAspPheGlyLeuSerLysMetGluGlnAsnGlyIleMetSerThrAlaCys 180
 Db 550 ATGATCACTGACTTGGTCTCTCCAGATGGAACAGATGGCATCATGCTCCATGCTGCTGT 609
 QY 181 GlyThrProGlyTyrValAlaProGluValLeuAlaGlnLysProTyrSerLysAlaVal 200
 Db 610 GGGACCCAGGCTAGTGGCTCCAGAAAGTGTGCTGCCCGAGAAACCTTACAGCAAGGCTGTG 669
 QY 201 AspCysTyrSerIleGlyValIleThrTyrIleLeuLeuCysGlyTyrProProPheTyr 220
 Db 670 GATTCTGCTCCATCGCGCTCATCCTCATATATGCTCTGTGGATACCCCGCTTCTAT 729
 QY 221 GluGluThrGluSerLysLeuPheGluLysIleLysGluGlyTyrTyrGluPheGluSer 240
 Db 730 GAAGAAACCGAGTCTTAAGCTTTTTCGAGAAGATCAAGAGGCGCTACTATGATGTTGAGTCT 789
 QY 241 ProPheTyrAspAspIleSerGluSerAlaLysAspPheIleCysHisLeuLeuGluLys 260
 Db 790 CCATTTCTGGATGACATTTTCTGAGTCAGCCCAAGGACTTTATTTGCCACTTGTCTTGAGAAG 849
 QY 261 AspProAsnGluArgTyrThrCysGluLysAlaLeuSerHisProThrIleAspGlyAsn 280
 Db 850 GATCCGACGAGCGGTACACCTGTGAGAAGGCCCTTGTAGTCATCCCTGGATGACGGAAC 909
 QY 281 ThrAlaLeuHisArgAspIleTyrProSerValSerLeuGlnIleGlnLysAsnPheAla 300
 Db 910 ACGGCCCTCCACCGGAGATCATCTACCCATCAGCTAGCCCTCCAGATCCAGAGAACTTTGCT 969
 QY 301 LysSerLysTrpArgGlnAlaPheAsnAlaAlaValValHisHisMetArgLysLeu 320

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 10:03:49 ; Search time 566 Seconds
(without alignments)
9736.113 Million cell updates/sec

Title: US-09-960-643-1
Perfect score: 2447
Sequence: 1 tggagtgaggagctcaagcag.....ttttctctataaaaaaaa 2447

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
N_Geneseq_101002.*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*		
24:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*		
SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	2447	100.0	2447 24 AAD36140 Human calmodulin k
2	1915	78.3	1956 22 AAI60703 Human polynucleoti
3	1854.6	75.8	2165 22 AAI58917 Human polynucleoti
4	1761.6	72.0	2689 22 AAS31014 Human diagnostic a
5	504.6	20.6	1565 24 ABL60905 Human Cam kinase I
6	504.6	20.6	1661 22 AAD18926 Human kinase (PKIN
7	504.6	20.6	1671 22 AAD25120 Nucleotide sequenc
8	504.6	20.6	1733 22 AAI58638 Human polynucleoti
9	503	20.6	1736 22 AAD18817 Human kinase (PKIN

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

10	503	20.6	2164	22	AAI60424	Human polynucleoti
11	502.2	20.5	1074	22	AA506708	Polynucleotide seq
12	502.2	20.5	1074	22	AAH25119	Nucleotide sequenc
13	502.2	20.5	1158	22	AAH25118	Nucleotide sequenc
14	497	20.3	1578	22	AAD04775	Human death domain
15	492.6	20.1	1694	22	AAC90432	Murine Dendritic C
16	471.4	19.3	1728	21	AAZ46162	cDNA sequence enco
17	428.8	17.5	1458	19	AAV18867	Calmodulin-depende
18	427	17.4	1032	22	AA506711	Polynucleotide seq
19	426.2	17.4	588	23	ABK43480	DNA encoding novel
20	418	17.1	1282	20	AAK06835	Disease associated
21	393.8	16.1	1512	23	AA570157	DNA encoding novel
22	337.2	13.8	362	20	AAK51838	Human secreted pro
23	291	11.9	476	22	AA527154	cDNA encoding nove
24	291	11.9	476	23	ABK43797	DNA encoding novel
25	290	11.9	2264	23	ABL02719	Drosophila melanog
26	286.4	11.7	1572	22	AAH99632	Human protein enco
27	285.6	11.7	1774	23	AA568767	DNA encoding novel
28	285.6	11.7	1774	23	AA590919	DNA encoding novel
29	245.2	10.0	2897	22	AAK30481	Human protein phos
30	245.2	10.0	2905	22	AAK94325	Human full-length
31	245.2	10.0	3017	22	AAI18832	Human kinase (PKIN
32	243.6	10.0	2297	22	AAH78261	Nucleotide sequenc
33	243.6	10.0	2747	22	AAI58143	Human polynucleoti
34	243.6	10.0	2840	22	AAI58144	Human polynucleoti
35	243.6	10.0	3124	24	ABA95682	Human protein kina
36	242	9.9	2494	22	AAK94614	Human full-length
37	242	9.9	3022	22	AAK94299	DNA encoding novel
38	217.8	8.9	709	23	ABK43538	Human cDNA encodin
39	215.4	8.8	1181	22	AA534308	cDNA encoding nove
40	215.4	8.8	1181	22	AA527166	DNA encoding novel
41	215.4	8.8	1181	23	ABK43848	PSKH-1 cDNA. Homo
42	197.2	8.1	3471	18	AAK71761	DNA encoding novel
43	194	7.9	1386	22	AAK52586	Human polynucleoti
44	194	7.9	1450	22	AAK51602	Human polynucleoti
45	185.2	7.6	852	21	AAC76557	Human OREX ORF2112

ALIGNMENTS

RESULT 1

AD36140	
ID	AD36140 standard; DNA; 2447 BP.
AC	AD36140;
XX	
DT	09-AUG-2002 (first entry)
XX	
DE	Human calmodulin kinase, CAMK-X1 gene.
XX	
KW	Human; cytosolic; antisense gene therapy; screening; protein kinase;
KW	cancer; liver; colon; tumour; inflammation; arthritic synovium; CAMK-X1;
KW	calmodulin kinase; enzyme; gene; chromosome Iq32.1-32.3; ds.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	70..1500
FT	/*tag= a
XX	/product= "Human CAMK-X1 protein"
PN	WO200224947-A2.
XX	
PD	28-MAR-2002.
XX	
PF	20-SEP-2001; 2001WO-IB02237.
XX	
PR	20-SEP-2000; 2000US-233999P.
PR	02-OCT-2000; 2000US-237419P.
PR	02-OCT-2000; 2000US-237423P.
PR	04-OCT-2000; 2000US-238558P.
PR	10-MAY-2001; 2001US-290555P.

XX (KINE-) KINETEK PHARM INC.
PA (UVBR-) UNIV BRITISH COLUMBIA.
XX
PI Yoganathan T, Delaney AD;
XX
DR WPI; 2002-394145/42.
DR P-PSDB; NAE22764.
XX
PT Diagnosing cancer, comprises determining the upregulation of expression
PT of a nucleic acid sequence encoding a protein kinase or upregulation of
PT expression of the protein kinase, in the cancer -
XX
PS Claim 16; Page 62-64; 87pp; English.
XX
CC The invention relates to a method for screening biologically active agent
CC that modulates cancer associated protein kinase function. The invention
CC also relates to a method for diagnosing cancer comprising determining the
CC upregulation of expression of a nucleic acid sequence encoding a protein
CC kinase. The method is useful for diagnosing cancer. A protein kinase is
CC useful for screening biological agents that modulate cancer associated
CC protein kinase function. Downregulating the activity of protein kinase is
CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon
CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy
CC derived tumours and inflammatory samples such as arthritic synovium, for
CC amplified DNA in the cell or increased expression of corresponding mRNA
CC or protein and is also useful to detect differences in expression levels
CC such as molecular weight, amino acid and nucleotide sequences between the
CC two cells. The present sequence is human calmodulin kinase CAMK-X1 gene
CC located on chromosome 1q32.1-32.3.
XX
SQ Sequence 2447 BP; 590 A; 707 C; 604 G; 546 T; 0 other;

Query Match
Best Local Similarity 100.0%; Score 2447; DB 24; Length 2447;
Matches 2447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGAGTGGGAGCTCAAGCAGGATTCTCCCGAGTCCCTGGCATCCTCAGAAAGCTTCAACT 60
Db 1 TGGAGTGGGAGCTCAAGCAGGATTCTCCCGAGTCCCTGGCATCCTCAGAAAGCTTCAACT 60

QY 61 CTGGAGGCAATGGGTGCGAAGGAAGAAGATGACTGCAGTCTCTGGAAGAACACACCACC 120
Db 61 CTGGAGGCAATGGGTGCGAAGGAAGAAGATGACTGCAGTCTCTGGAAGAACACACCACC 120

QY 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAAGTT 180
Db 121 AACATCCGGAACCTTCATTTTATGGAAGTCTGGGATCAGAGCTTCTCAGAAAGTT 180

QY 181 TTCCTGGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 240
Db 181 TTCCTGGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTCA 240

QY 241 CCTGCTTCCGGACAGCAGCTCGAGATGAGATTGCTGTTGGAAGAAAGATCAAGCAT 300
Db 241 CCTGCTTCCGGACAGCAGCTCGAGATGAGATTGCTGTTGGAAGAAAGATCAAGCAT 300

QY 301 GAAACATTTGTGACCTCGAGGACATCTATGAGAGCACCACCCTACTACTGTCATG 360
Db 301 GAAACATTTGTGACCTCGAGGACATCTATGAGAGCACCACCCTACTACTGTCATG 360

QY 361 CAGCTTGTTCCTGGTGGGAGCTTTGACCGGATCCTTGGAGCGGGGTGTCTACACAGAG 420
Db 361 CAGCTTGTTCCTGGTGGGAGCTTTGACCGGATCCTTGGAGCGGGGTGTCTACACAGAG 420

QY 421 AAGGATGCCAGTCTGGTATCCAGAGGTCTTTGTCGGCAGTGAATACCTACATGAGAAAT 480
Db 421 AAGGATGCCAGTCTGGTATCCAGAGGTCTTTGTCGGCAGTGAATACCTACATGAGAAAT 480

QY 481 GGCATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTTACCCCTGAAGAGAAC 540
Db 481 GGCATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTTACCCCTGAAGAGAAC 540

QY 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCTCAAGATGGAACAGAAATGSCATCATGTCC 600
Db 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCTCAAGATGGAACAGAAATGSCATCATGTCC 600

QY 601 ACTGCTGTGGGACCCCGAGGCTACGTGGCTCCAGAGTGTGCTGGCCAGAAACCTTACAGC 660
Db 601 ACTGCTGTGGGACCCCGAGGCTACGTGGCTCCAGAGTGTGCTGGCCAGAAACCTTACAGC 660

QY 661 AAGGCTGTGATGTCTGCTGCTCCATCGGCGTCACTCACCTTACATATTGCTCTGTGTGATACCCC 720
Db 661 AAGGCTGTGATGTCTGCTGCTCCATCGGCGTCACTCACCTTACATATTGCTCTGTGTGATACCCC 720

QY 721 CCGTCTTATGAAGAACGAGTCTAAGCTTTTTCGAGAAGATCAAGGAGGCTACTATGAG 780
Db 721 CCGTCTTATGAAGAACGAGTCTAAGCTTTTTCGAGAAGATCAAGGAGGCTACTATGAG 780

QY 781 TTTGAGTCTCCATTCTGGGATGACATTTCAGTACAGCCAGGACTTTATTTTCCCACTTG 840
Db 781 TTTGAGTCTCCATTCTGGGATGACATTTCAGTACAGCCAGGACTTTATTTGCCACTTG 840

QY 841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTTGAGTCAATCCCTGGATT 900
Db 841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTTGAGTCAATCCCTGGATT 900

QY 901 GACGGAACAGGCGCTCCACCGGGACATCTACCCATCAGTACGCTCCAGATCCAGAAAG 960
Db 901 GACGGAACAGGCGCTCCACCGGGACATCTACCCATCAGTACGCTCCAGATCCAGAAAG 960

QY 961 AACTTTCTTAAGAGCAAGTGGAGCAAGCTTCAACGACGACGCTGTGTGTGACCAACATG 1020
Db 961 AACTTTCTTAAGAGCAAGTGGAGCAAGCTTCAACGACGACGCTGTGTGTGACCAACATG 1020

QY 1021 AGGAAGCTACACATGAACCTGCACAGCCCGGGGTCCGCCAGAGGTGGAGAACAGGCCG 1080
Db 1021 AGGAAGCTACACATGAACCTGCACAGCCCGGGGTCCGCCAGAGGTGGAGAACAGGCCG 1080

QY 1081 CTTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCAACATCAACCGAG 1140
Db 1081 CTTGAAACTCAAGCTCAGAAACCTCTAGACCCAGCTCCCTGAGATCAACATCAACCGAG 1140

QY 1141 GCACCTCTCTGGACACAGTGTAGCAGTCCCTGCCTGACCCAAATTAACCTGCCAGCAT 1200
Db 1141 GCACCTCTCTGGACACAGTGTAGCAGTCCCTGCCTGACCCAAATTAACCTGCCAGCAT 1200

QY 1201 GGGCGCCGGCCCACTGCCCCCTGTGGCAGGTCCCTCAACTGCTGTGTCATGCTCCCTC 1260
Db 1201 GGGCGCCGGCCCACTGCCCCCTGTGGCAGGTCCCTCAACTGCTGTGTCATGCTCCCTC 1260

QY 1261 CACATCAGCAGCAGCTGGTGCCCATGCATCAGGGTCCCTGGCGCCGGCCCTGTGGC 1320
Db 1261 CACATCAGCAGCAGCTGGTGCCCATGCATCAGGGTCCCTGGCGCCGGCCCTGTGGC 1320

QY 1321 TGTGCTCCAGCTGCCCTGAACATTTGGGAGCAAAAGAAAGTCCCTACTGCTCTGAGCCC 1380
Db 1321 TGTGCTCCAGCTGCCCTGAACATTTGGGAGCAAAAGAAAGTCCCTACTGCTCTGAGCCC 1380

QY 1381 ACATCTCTAAAAGGCGCAACAAAAACAGAACTTCAAGTGGAGGTCAATGGTACAGATT 1440
Db 1381 ACATCTCTAAAAGGCGCAACAAAAACAGAACTTCAAGTGGAGGTCAATGGTACAGATT 1440

QY 1441 AAAGCCAGTGGAGTCCCATGCGGCGAGGCGAGCTGGAGTCTGTCTCATATTATGTGA 1500
Db 1441 AAAGCCAGTGGAGTCCCATGCGGCGAGGCGAGCTGGAGTCTGTCTCATATTATGTGA 1500

QY 1501 TTCTGAGGAGCTGTGCTTATGTCAATTTTTCAGGAGACATATTCAACTCCTCTGTCT 1560
Db 1501 TTCTGAGGAGCTGTGCTTATGTCAATTTTTCAGGAGACATATTCAACTCCTCTGTCT 1560

QY 1561 CTTCCAAACCTGTGTCTATCCGGCAGGAGGAGGAGGAGCAAGTGGAGGAGGCGCTT 1620
Db 1561 CTTCCAAACCTGTGTCTATCCGGCAGGAGGAGGAGGAGCAAGTGGAGGAGGCGCTT 1620

QY 1621 AGCAGGAGCAGTTTCTGGCCAGAAAGCAGCAGCTGTGTCAGCGGGGAGCCCTCATAG 1680

|||||
Db 1621 AGCAGGAGCAGTTCTGGCCAGAACACACAGCCTGTGCGAGGGGCGCCCTCATAG 1680
Qy 1681 GAGGCCAGGAGGAGGCCAGCGCTAGAACGCTTGTGAAGCTGTGAGCAGGAGAACG 1740
Db 1681 GAGGCCAGGAGGAGGCCAGCGCTAGAACGCTTGTGAAGCTGTGAGCAGGAGAACG 1740
Qy 1741 GGTGCCACACAGCTTCCAGTCTCCCTGACCTGCCTGCTATGCCCCACACCTACGTG 1800
Db 1741 GGTGCCACACAGCTTCCAGTCTCCCTGACCTGCCTGCTATGCCCCACACCTACGTG 1800
Qy 1801 CCGTGGCTCTGTGCACTGACGTAGTAGCTCTCGCTGGCTGTGCTGCTTGTGCTGTA 1860
Db 1801 CCGTGGCTCTGTGCACTGACGTAGTAGCTCTCGCTGGCTGTGCTGCTTGTGCTGTA 1860
Qy 1861 AAAGCTTAATGGGCTGCCAGGTGTGTACCTTCTCCAGCAAAAGCCATATGGAGCATC 1920
Db 1861 AAAGCTTAATGGGCTGCCAGGTGTGTACCTTCTCCAGCAAAAGCCATATGGAGCATC 1920
Qy 1921 TACCCAGACTCCCACTCTGCACACACTCCTCCACCTCTCAAGCCTCCCAACCTCTGGC 1980
Db 1921 TACCCAGACTCCCACTCTGCACACACTCCTCCACCTCTCAAGCCTCCCAACCTCTGGC 1980
Qy 1981 CAGATTGGGCTCAATTAATGCTGCTGCCATCTGCATGAATGACAGGAGCTCCCA 2040
Db 1981 CAGATTGGGCTCAATTAATGCTGCTGCCATCTGCATGAATGACAGGAGCTCCCA 2040
Qy 2041 TGGTGTCTGCTGTGAGCTTCAAGTTCTAATCTTAATCTTAATCTTAATCTTAATCT 2100
Db 2041 TGGTGTCTGCTGTGAGCTTCAAGTTCTAATCTTAATCTTAATCTTAATCTTAATCT 2100
Qy 2101 TGGCTGAGACCCAGCAGCAGCTCTGCGCTCTGCTGCTCAATCTCAATCTCAATCT 2160
Db 2101 TGGCTGAGACCCAGCAGCAGCTCTGCGCTCTGCTGCTCAATCTCAATCTCAATCT 2160
Qy 2161 GCCACACCCCTCCAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2220
Db 2161 GCCACACCCCTCCAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2220
Qy 2221 ATCTGCTGCCACATCTCTCCGTCACACACCCATGAGTTACCTTGGAGTGAAGTGA 2280
Db 2221 ATCTGCTGCCACATCTCTCCGTCACACACCCATGAGTTACCTTGGAGTGAAGTGA 2280
Qy 2281 TTTAATGTCTGCCAGGATCTAATCTGCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCT 2340
Db 2281 TTTAATGTCTGCCAGGATCTAATCTGCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCT 2340
Qy 2341 AGCACACATCTGTCTCCAGTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db 2341 AGCACACATCTGTCTCCAGTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Qy 2401 AGTGTGTGCTTAATAAATGGACATATTTTCTCTAAATAAATAAATAAATAAATAA 2447
Db 2401 AGTGTGTGCTTAATAAATGGACATATTTTCTCTAAATAAATAAATAAATAAATAA 2447

RESULT 2
AAI60703
ID AAI60703 standard; cdna; 1956 BP.
XX
AC AAI60703;
XX
DT 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 4692.
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX

OS Homo sapiens.
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB; AAM41547.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4692; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1956 BP; 474 A; 559 C; 508 G; 414 T; 1 other;

Query Match 78.3%; Score 1915; DB 22; Length 1956;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1941; Conservative 0; Mismatches 6; Indels 3; Gaps 2;
Qy 12 CTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTCTCAGAAGCTTCAACTCTGGAGCAAT 71
Db 7 CTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTCTCAGAAGCTTCAACTCTGGAGCAAT 66
Qy 72 GGGTCGAAGAGAGATGACTGCAGTTCCTGGAGAAACAGACCACACATCCGGAA 131
Db 67 GGGTCGAAGAGAGATGACTGCAGTTCCTGGAGAAACAGACCACACATCCGGAA 126
Qy 132 AACCTTCAATTTTATGGAAGTGGGATCAGGAGCTTCTCAGAAGTTTCTCTGGTGA 191
Db 127 AACCTTCAATTTTATGGAAGTGGGATCAGGAGCTTCTCAGAAGTTTCTCTGGTGA 186
Qy 192 GCAGAACTGACTGGGAAGCTTTTGTCTCTGAAGTGCATCAAGAGTCACCTGCTTCCG 251
Db 187 GCAGAACTGACTGGGAAGCTTTTGTCTCTGAAGTGCATCAAGAGTCACCTGCTTCCG 246
Qy 252 GCAGAGCAGCTGGGAAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAACATTGT 311
Db 247 GCAGAGCAGCTGGGAAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGAAACATTGT 306

Qy	312	GACCC	TGGAGGACATCTATGAGAGCACCACCACCTACTACTGTCATGCGAGCTTGTTTC	371
Db	307	GACCC	TGGAGGACATCTATGAGAGCACCACCACCTACTACTGTCATGCGAGCTTGTTTC	366
Qy	372	TGCTG	GGGAGCTCTTTTGACCGGATCCTGGAGGGGTGTCTACACAGAGAAGATGCCAG	431
Db	367	TGCTG	GGGAGCTCTTTTGACCGGATCCTGGAGGGGTGTCTACACAGAGAAGATGCCAG	426
Qy	432	TCGTG	GTATCCACGAGTCTTGTCGGCAGTGAATACCTACATCAGAAATGGCATCGTCCA	491
Db	427	TCGTG	TATCCACGAGTCTTGTCGGCAGTGAATACCTACATCAGAAATGGCATCGTCCA	486
Qy	492	CAGAG	ACTTAAAGCCCGAAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTTAAGATCAT	551
Db	487	CAGAG	ACTTAAAGCCCGAAAACCTGCTTTACCTTACCCCTGAAGAGAACTCTTAAGATCAT	546
Qy	552	GATCA	CTGACTTTGGTCTCTCCAAGATGGACAGAAATGGCATCATGTCCACTGCCTGTGG	611
Db	547	GATCA	CTGACTTTGGTCTCTCCAAGATGGACAGAAATGGCATCATGTCCACTGCCTGTGG	606
Qy	612	GAGCC	CAGGCTACGTGGCTCCAGAAAGTGTGTCGCCAGAAACCTTACAGCAAGGCTGTGGA	671
Db	607	GAGCC	CAGGCTACGTGGCTCCAGAAAGTGTGTCGCCAGAAACCTTACAGCAAGGCTGTGGA	666
Qy	672	TTGCT	GGTCCATCGGGCTCATCACCTACATATTGCTCTGTGGATACCCCGCTTCATGA	731
Db	667	TTGCT	GGTCCATCGGGCTCATCACCTACATATTGCTCTGTGGATACCCCGCTTCATGA	726
Qy	732	AGAA	CGGAGTCTTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGTTTGAGTCTCC	791
Db	727	AGAA	CGGAGTCTTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGTTTGAGTCTCC	786
Qy	792	ATTCT	GGGATGACATTTCTGATCTAGCCAAAGACTTTATTGGCACTTTGCTTGAGAAGGA	851
Db	787	ATTCT	GGGATGACATTTCTGATCTAGCCAAAGACTTTATTGGCACTTTGCTTGAGAAGGA	846
Qy	852	TCCGA	CAGAGCGGTACACCTGTGAGAAGCCTTTGAGTCATCCCTGGATTGACGGAAACAC	911
Db	847	TCCGA	CAGAGCGGTACACCTGTGAGAAGCCTTTGAGTCATCCCTGGATTGACGGAAACAC	906
Qy	912	GGCC	TCCACCGGGACATCTACCCATCAGTCAGCCTCCAGATCCAGAAAGCTTTGCTAA	971
Db	907	AGCC	TCCACCGGGACATCTACCCATCAGTCAGCCTCCAGATCCAGAAAGCTTTGCTAA	966
Qy	972	GAGCA	GTTGGAGCAAGCCTTCAACGACGACGCTGTGTGTGCACACATGAGGAAGCTACA	1031
Db	967	GAGCA	GTTGGAGCAAGCCTTCAACGACGACGCTGTGTGTGCACACATGAGGAAGCTACA	1026
Qy	1032	CATGA	ACCTGCAACGCGCGGCTCCGCCACGAGGTGGAGAACAGGCGCCTGAAACTCA	1091
Db	1027	CATGA	ACCTGCAACGCGCGGCTCCGCCACGAGGTGGAGAACAGGCGCCTGAAACTCA	1086
Qy	1092	AGCC	TAGAAACCTCTAGACCCAGCTCCCTTGAGATCACATACCGAGGCACTGTGCTT	1151
Db	1087	AGCC	TAGAAACCTCTAGACCCAGCTCCCTTGAGATCACATACCGAGGCACTGTGCTT	1146
Qy	1152	GGACC	CAGTGTAGCACTCCCTGCCCTGACCCCAATTACCCCTGCCAGATGGCCCGCGGCC	1211
Db	1147	GGACC	CAGTGTAGCACTCCCTGCCCTGACCCCAATTACCCCTGCCAGATGGCCCGCGGCC	1206
Qy	1212	CAGT	GCCCTGTGGCAGGTCCCTCACTGCTGTGTCAATGGCTCCCTCCACATCAGCAG	1271
Db	1207	CAGT	GCCCTGTGGCAGGTCCCTCAACTGCTGTGTCAATGGCTCCCTCCACATCAGCAG	1266
Qy	1272	CAGC	TGTCCTCATCATCAGGGGTCCCTGTGCCCGCGGCCCTGTGTGCTGTGCTCCAG	1331
Db	1267	CAGC	TGTCCTCATCATCAGGGGTCCCTGTGCCCGCGGCCCTGTGTGCTGTGCTCCAG	1326
Qy	1332	CTGC	TGACATTTGGGAGCAAAAGATTCCTTCTACTGTCTGTAGGCCCACTCCTCAA	1391
Db	1327	CTGC	TGACATTTGGGAGCAAAAGATTCCTTCTACTGTCTGTAGGCCCACTCCTCAA	1386

Qy	1392	AAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTGATGGTACCAGTTAAAGCCAGTGG	1451
Db	1387	AAAGGCCAACAAAAACAGAACTTCAAGTCGGAGGTGATGGTACCAGTTAAAGCCAGTGG	1446
Qy	1452	CAGCTCCCAGTCCCGGCAGGGCAGACTCGAGCTGCTCATTAATGATGATTCCCTGGAGCC	1511
Db	1447	CAGCTCCCAGTCCCGGCAGGGCAGACTCGAGCTGCTCATTAATGATGATTCCCTGGAGCC	1506
Qy	1512	TGTGCCCTATGTCACTGCAATTTTTCAGGAGACATATTCAACTCCTCTGCTCTTCCAAACCT	1571
Db	1507	TGTGCCCTATGTCACTGCAATTTTTCAGGAGACATATTCAACTCCTCTGCTCTTCCAAACCT	1566
Qy	1572	GGTGCTCTATCCGGCAGAGGGAAGGCAGAGCAAGTGGAGCAGGGCTTAGCAGAGCAG	1631
Db	1567	GGTGCTCTATCCGGCAGAGGGAAGGCAGAGCAAGTGGAGCAGGGCTTAGCAGAGCAG	1626
Qy	1632	TTTTCTGGCCAGAGCACACCAGCTGCTGCCAGCGGGCAGCCCTCATAGGAGGCCACAGA	1691
Db	1627	TTTTCTGGCCAGAGCACACCAGCTGCTGCCAGCGGGCAGCCCTCATAGGAGGCCACAGA	1686
Qy	1692	GGGAGCCCCAAGCGGTAGAAGCCCTTGTTGAAGCTGTGAGCAGGAGAAGCGGTGCCACCA	1751
Db	1687	GGGAGCCCCAAGCGGTAGAAGCCCTTGTTGAAGCTGTGAGCAGGAGAAGCGGTGCCACCA	1746
Qy	1752	GCTTCCAGGTCTCCCTGACCTGCCCTGTATATGCCCCACACCCCTACGT-GCGGTGGCTCT	1810
Db	1747	GCTTCCAGGTCTCCCTGACCTGCCCTGTATATGCCCCACACCCCTACGTGGCGTGGCTCT	1806
Qy	1811	GTGCAGTGT--ACGTAGATAGCTCTCGCTGGGTCTGTGCTGTGTTGCTGTGAANAAGCTTA	1868
Db	1807	GTGCAGTGTACGTTAGATAGCTCTCGCTGGGTCTGTGCTGTGTTGCTGTGAANAAGCTTA	1866
Qy	1869	ATGGGCTGGCCAGGCTGTGTACCTTTCTCCAAGCAAAGGCATATATGGAGCATCTACCCAGA	1928
Db	1867	ATGGGCTGGCCAGGCTGTGTACCTTTCTCCAAGCAAAGGCATATATGGAGCATCTACCCAGA	1926
Qy	1929	CTCCCACTCTGCACACACTCACTCCCACT	1958
Db	1927	CTCCCACTCTGNACACACTCACTCCCACT	1956

RESULT 3

AAI58917
ID AAI58917 standard; cDNA; 2165 BP.

XX
AC

DT 22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 1120.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.

XX Homo sapiens. OS

XX PN WO200153312-A1.

XX
PD
26-JUL.-2001

XX
PF 26-DEC-2000: 2000WO-IIS34263-

XX
PR 27-JAN-2000: 2000US-0488725

PR 25-APR-2000; 2000US-0552317.
PR 00-JUL-2000; 2000US-0598042

PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450

PR 14-SEP-2000; 2000US-0662191.

Db 386 TCTGAAGTGCATCAAGAAAGTCACTGCCTTCCGGGACAGCGCTGGAGAAATGAGATTGC 445
Qy 279 TGTGTTGAAAAGATCAAGCATGAAAACATTTGTGACCCCTGGAGGACATCTATGAGAGCAC 338
Db 446 TGTGTTGAAAAGATCAAGCATGAAAACATTTGTGACCCCTGGAGGACATCTATGAGAGCAC 505
Qy 339 CACCCACTACTACTCGTGCATGCAGCTTGTCTTGTGGGGAGCTCTTTGACCGGATCCT 398
Db 506 CACCCACTACTACTCGTGCATGCAGCTTGTCTTGTGGGGAGCTCTTTGACCGGATCCT 565
Qy 399 GGAGCGGGTGTCTACACAGAGAAGGATGCCAGTCTGTGTGATCCACAGAGTCTTGTGCGC 458
Db 566 GGAGCGGGTGTCTACACAGAGAAGGATGCCAGTCTGTGTGATCCACAGAGTCTTGTGCGC 625
Qy 459 AGTGAATACCTACATGAGAAAGGATGCCAGTCTGTGTGATCCACAGAGTCTTGTGCGC 518
Db 626 AGTGAATACCTACATGAGAAAGGATGCCAGTCTGTGTGATCCACAGAGTCTTGTGCGC 685
Qy 519 TTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTTTGGTCTGTCCAAGAT 578
Db 686 TTACCTTACCCCTGAAGAGAACTCTAAGATCATGATCACTGACTTTGGTCTGTCCAAGAT 745
Qy 579 GGAACGAATGGCATCATGTCCACTGCCTGTGGGACCCAGAGTCTACGTGGCTCCAGAGT 638
Db 746 GGAACGAATGGCATCATGTCCACTGCCTGTGGGACCCAGAGTCTACGTGGCTCCAGAGT 805
Qy 639 GCTGGCCAGAAACCTACAGAAAGCTGTGATTTGCTGGTCCATCGGGGTCAACCTA 698
Db 806 GCTGGCCAGAAACCTACAGAAAGCTGTGATTTGCTGGTCCATCGGGGTCAACCTA 865
Qy 699 CATATTGCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAA 758
Db 866 CATATTGCTGTGGATACCCCGCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAA 925
Qy 759 GATCAAGAGGGCTACTAGATTGAGTCTCCATTTCTGGGATGACATTTCTGAGTACG 818
Db 926 GATCAAGAGGGCTACTAGATTGAGTCTCCATTTCTGGGATGACATTTCTGAGTACG 985
Qy 819 CAAGGACTTTATTGGCACTTGTCTGAGAAGGATCCGAACGA - GGGGTACACCTGTGAGA 877
Db 986 CAAGGACTTTATTGGCACTTGTCTGAGAAGGATCCGAACGAAGGGGTAAACACTGTGAGA 1045
Qy 878 AGGCCTTGAATCATCCCTGGATTGACGAAACAGCGCCCTCCACCGGAGACATCTACCCAT 937
Db 1046 AGGCCTTGAATCATCCCTGGATTGACGAAACAGCGCCCTCCACCGGAGACATCTACCCAT 1105
Qy 938 CAGTCA - GCCTCAGATCCAGAAAGACTTTGCTAGAGAAAGTGGAGGCAAGCCCTCAAC 996
Db 1106 CAGTCAAGGGCTCAGATCCAGAGAACTTTTGCTAGAGCAAGTGGAGGCAAGCCCTCAAC 1165
Qy 997 GCAGCAGCTGTGT - GCACACATGAGAAAGTACACATGAACCTGCACA - GCCTGGGG 1054
Db 1166 GCAGCAGCTGTGTAGCACAAATGAGAAAGTACACATGAACCTGCACACCGCCGGGG 1225
Qy 1055 TCCGCCCAGAGGTGGAGAACAGSCGCCCTGAAACTCAAGCCCTCAGAAACCTTAGACCCA 1114
Db 1226 TCCGCCCAGAGGTGGAGAACAGSCGCCCTGAAACTCAAGCCCTCAGAAACCTTAGACCCA 1285
Qy 1115 GCTCCCTTGATATCACCATCAGGAGGACCTGTCTGACACACAGTGTAGACATCCCTG 1174
Db 1286 GCTCCCTTGATATCACCATCAGGAGGACCTGTCTGAGACACAGTGTAGACATCCCTG 1345
Qy 1175 CCCTGACCCAAATTA - - - CCCTGACAGATGGCGCGGCCACTGCCCTGTGGCAGGT 1231
Db 1346 CCCTGACCCAAATATACCTTGCCAGCATCGCGCGGCCCACTGCCCTGTGGCAGGT 1405
Qy 1232 CCCTCAACTGCCTGGTCAATGGTCCCTCCACATCAGCAGAGCCGTGGTCCCATGCATC 1291
Db 1406 CCCTCAACTGCCTGGTCAATGGTCCCTCCACATCAGCAGAGCCGTGGTCCCATGCATC 1465
Qy 1292 AGGGGTCCTTGGCGCGGCCCTGTGGTGTGCTGTCTCC - AGCTGCTGAACATTTGGGAGC 1350
Db 1466 AGGGGTCCTTGGCGCGGCCCTGTGGTGTGCTGTCTCCAAAGCTGCTTGAACATTTGGGAGC 1525

Qy 1351 AAAGGAAGTCCCTCTACTGCTGTGAGCCACACTCTCTCAAAAGGCGCAACAAAACAG 1410
Db 1526 AAAGGAAGTCCCTCTACTGCTGTGAGCCACACTCTCTCAAAAGGCGCAACAAAACAG 1585
Qy 1411 AACTTCAAGTCCGAGGTCATGGTACCAGTTAAAGCCAGTGGCA - GCTCCCACTGCCGGC 1469
Db 1586 AACTTCAAGTCCGAGGTCATGGTACCAGTTAAAGCCAGTGGCATGCTTCCACTGCCGGC 1645
Qy 1470 AGGCAGACTGGAGTCTGTCTCATTTATGTGATTCCTGGAGCCTGTGCTATGTCAC - -TG 1527
Db 1646 AGGCAGACTGGAGTCTGTCTCATTTATGTGATTCCTGGAGCCTGTGCTATGTCACGTG 1705
Qy 1528 CAATTTTCAGAGACATATTCAACTCTCTGCTCTTCCAAACCTGGTGTCTATCCGGCAG 1587
Db 1706 AATTTTCAGAGACATATTCAACTCTCTGCTCTCCAAACCTGGTGTCTATCCGGCAG 1765
Qy 1588 AGGAGGAAGCAGAGCAAGTGGAGCAGGGCTTAGCAGGACAGTTTCT - GGCCAGAAGC 1646
Db 1766 AGGAGGAAGCAGAGCAAGTGGAGCAGGGCTTAGCAGGACAGTGTCTGGCCAGAAGC 1825
Qy 1647 ACCA - - - GCCTGCTGCCAGCGGGCAGCCCTCATAGGAGGCCAGGAGGAGGCCCAA 1702
Db 1826 ACCATGCTGCTTCCAGCTGGGGCAGCCCGCTCATAGGAGGCCAGGAGGAGGCCCAA 1885
Qy 1703 GGCCTAGAAAGCCTT - GTTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAG - - 1759
Db 1886 GGCCTAGAAAGCCTTGGTGTGAAGCTGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGT 1945
Qy 1760 -GTCTCCCTGACCTGCTCTATGCCACACACCTTAC - -GTGCCCTGGCTGTGTCAG 1816
Db 1946 TCTCCGTGACCTGCGCTGCTATGCCACACACCTTACAGTGGCCGTGGCTGTGTCAG 2005
Qy 1817 TGT - -AGCTAGATAGCTCTCG - COTGGGTCTGTGCTGTTTGTGCTGTAAGAGCTTAA -TG 1871
Db 2006 TGTTACGTTAGATAGCTCTCGCCCTGGGTCTGTGCTGTTTGTGCTGTAAGAGCTTAAATGG 2065
Qy 1872 GGCCTGGCAGCTGTGTCAACCTTCTCAAGCAAAAGCCATAT - - - GGAGATCTTACCC - - 1925
Db 2066 GCTGGCCAGGCTGTGTCAACCTTCTCAAGCAAAAGCCATATAGTGTGAGCATCTATCCAG 2125
Qy 1926 -AGACTCCCACTGTGCACACACTCACCTCCACCTCTCAAGCCTCCAAACCTCTTGGCCAGA 1984
Db 2126 AGATCCCACTGTGTGCACACTCTCTCACTCACTCTAAAGTCTCAACCTCTTGGCCAGA 2185
Qy 1985 TTGGGCTCATTAAT - GTGCTGTGCTGCCCATCTG - CATGAATGACAGGAGCTCCCATG 2042
Db 2186 TTGGGCTCATTAATGGTGTGCTGCCCATCTGCCCATGAATACCCAGCTCCCATG 2245
Qy 2043 GTGCTGTGCT - GTGAGCTCTTCAAGTCTTAATCTTAACTTCCAGGATTAAGTCCCAAGT 2101
Db 2246 GTGCTGTGCTGTGAGCTCTTCAAGTCTTAATCTTAACTTCCAGGATTAAGTCCCAAGT 2305
Qy 2102 GCGTGTAGACCCAGCCA - - -GCACACTTGTGGCCCTTCTCCCTGCTCAATCTAAAAGCA 2158
Db 2306 GCGTGTAGAACCCAGCAAGCAGACACTTCTTGGCCCTTCTCCCTGCTCAATCTAAAAGCA 2365
Qy 2159 GTGCCACACCTCCAAA - GTTGAATAGAAAGTTTCATGAGTAAG - GGCCTGCAAGGAAT 2216
Db 2366 GTGCCACACCTCCAAAAGGTGGAATAGAAAGTTTCATGAGTAAGGCTGCAACGTTG 2425
Qy 2217 TCTTATCTGCGCCA - - - - -CATGCTCTCCGTGCACA - - - -CACCCAATGGAGTTA 2262
Db 2426 AACTACTTAGTCCATGAGTCCAGCACTGTCTCGTGCACAGCAGCCCAATGGACGTTA 2485
Qy 2263 ACCTTGAAGTGTAC - - - -TATTTTAAATGTGTGCCAGAG - -TTCATACTGCTGTG - - 2315
Db 2486 ACCTTGAAGTGTGATATATTAAGTGTGTGCCAGAGGTTCTAAGTCTGCTGCTGTG 2545
Qy 2316 TTCCCTTTTCTCTCTTCAA - - - - -AGTCCACACACCATTTCTTCTGCTTCCC - - - 2364
Db 2546 TCCGCTTTTCTCTCTTGAATGTGCCAGTGCAGATCCATATCTTGTGCTGTCCCGAGG 2605

KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW obesity; restorative therapy; immunomodulatory; vaccine; cardiovascular;
KW antimicrobial; cytostatic; antiinflammatory; asthma; ss.
XX
OS Homo sapiens.

PH Key Location/Qualifiers
FT CDS 159..1226
FT /*tag= a
FT /product= "Human PKIN-11 protein"

XX W0200181555-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12992.

XX 20-APR-2000; 2000US-199021P.

XX 28-APR-2000; 2000US-200226P.

XX 05-MAY-2000; 2000US-202339P.

XX 11-MAY-2000; 2000US-203505P.

XX 18-MAY-2000; 2000US-205564P.

XX 26-MAY-2000; 2000US-207739P.

XX 01-JUN-2000; 2000US-208795P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
PI Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;
PI Patterson C, Burrill JD, Marcus GA, Zingler KA, Recipon SA, Lu Y;
PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
PI Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;
PI Gururajan R;

XX WPI: 2001-611740/70.

DR P-PSDB; AAE11777.

XX Human kinases and nucleic acids, useful for preventing diagnosing and
PT treating cancers, inflammation and immune disorders -

XX Claim 5; Page 159; 166pp; English.

XX The present invention relates to human kinases (PKIN) and the nucleic
CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
CC used in the prevention, diagnosis and treatment of diseases cancers,
CC adenocarcinoma, leukemia, sarcoma, immune disorder, Addison's disease,
CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
CC gout, microbial infections, cardiovascular disease and/or inflammation,
CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC infarction, cataract, growth and development disorder, seizure disorder,
CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC PKIN may be used to treat disorders associated with decreased PKIN
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of PKIN by expressing inactive proteins or to
CC supplement the patients own production of PKIN. PKIN nucleic acids may be
CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
CC a host cell and culturing the cell to express the protein. PKIN nucleic
CC acid and its complementary sequences may also be used as DNA probes in
CC diagnostic assays to detect and quantitate the presence of similar
CC nucleic acid sequences in samples and therefore which patients may be
CC in need of restorative therapy. The present sequence is human PKIN-11
CC cDNA.

XX Sequence 1661 BP; 408 A; 452 C; 448 G; 353 T; 0 other;

XX Query Match 20.6%; Score 504.6; DB 22; Length 1661;
XX Best Local Similarity 71.3%; Pred. No. 4.4e-132;
XX Matches 695; Conservative 0; Mismatches 274; Indels 6; Gaps 2;
XX

XX

XX

XX

XX

XX

Db 155 GCCATGCGCCGGGAGAACGGGAGACAGAGCTCTCTCTGGAAAAAGCAAGCTGAAGACAT 214
QY 126 CCGGAAAAACCTTTCATTTTATGGAAGTCTCGGATCAGGAGCTTTCTCAGAAAGTTTTCCT 185
Db 215 CAAGAAGATCTTCGAGTTTCAAAGAGACCCCTCGGAACGGGGCCCTTTTCGGAAGTGGTTT 274
QY 186 GGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATC---AAGAAGTCACC 242
Db 275 AGCTGAAGAGAGGCAACTGGCAAGCTCTTTGCTGTGAAGTGTATCTCCTCAAGAGGCGCT 334
QY 243 TGCCTTCGCGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAAGATCAACAGCATGA 302
Db 335 GAAGGCAAGGAAGCAGCAGCATAGAGATGACATGCCGTCTCTGAGAAAGATTAGCATGA 394
QY 303 AAACATTGTGACCTCGGAGGACATCTATGAGAGACACCACCCTACTACTCTGGTCATGCA 362
Db 395 AAATATTGTTGCCCTGGAAGACATTTATGAAGCCCAATCACCCTGTACTTGGTCATGCA 454
QY 363 GCTTGTTTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAGAA 422
Db 455 GCTGTGTCCGGTGGAGAGCTGTTTGACCGGATAGTGGAGAAAGGGTTTTATACAGAGAA 514
QY 423 GGATGCCAGTCTGTGATCCAGAGCTTCTCGGCGAGTGAATACCTACTACATGAGAATGG 482
Db 515 GSATGCCAGCACTCTGATCCGCCAAGTCTTGGACGCCGTGTACTATCTCCACAGAATGG 574
QY 483 CATCTGCACAGAGACTTAAAGCCCGGAAACCTGCTTTACCTTACCTTCAAGAGAACTC 542
Db 575 CATCTGCACAGAGACCTCAAGCCCGGAAATCTCTGTACTACAGTCAAGATGAGGAGTC 634
QY 543 TAAGATCATGATCACTGACTTTTGGTCTCTCAAGATGGAACAGAATGG---CATCATGTC 599
Db 635 CAAAATAATGATCAGTGAATTTGGATTGTCAAAATGGAGGGCAAGAGATGTGATGTC 694
QY 600 CACTGCCCTGTGGACCCAGCTAGTGGTCTCCAGAAAGTGTGCGCCAGAAACCTACAG 659
Db 695 CACTGCCCTGTGAACTCCAGGCTATGTCGCTCCTGAACTCCTCGCCAGAAACCTTACAG 754
QY 660 CAAGCTGTGATGCTGCTCCATCGGCGTCACTACCATATGCTCTGTGGATACCC 719
Db 755 CAAGCCCTTGACTGCTGCTCCAGTGAATGCTGCTTACATCTGTCTCGGCTACCC 814
QY 720 CCGCTTCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGA 779
Db 815 TCCTTTTATGATGAATAGACTCCAAGCTCTTTGAGCAGATCCTCAAGCGGGAATATGA 874
QY 780 GTTGAGTCTCCATCTCGGATGACATTTCTGAGTACGCAAGGACTTTATTTGCCACTT 839
Db 875 GTTTGACTCTCCCTACTCGGATGACATCTCGGACTCTGCAAAAGACTTCATTCGGAACCT 934
QY 840 GCTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCCCTTGAGTCACTCCCTGGAT 899
Db 935 GATGAGAAGGACCGGATAAAGATACACGCTGTGAGCAGGAGCTCGGCACCATGGAT 994
QY 900 TGACGGAACACGGCCCTTCCACCGGAGCATCTACCCATAGTACGCTCCAGATCCAGAA 959
Db 995 CGCTGGTGACACAGCCCTTCAACAAAATCCACGAGTCCGTCAGCGGCCAGATCCGGAA 1054
QY 960 GAACCTTCTTGAAGCAGTGGAGCAAGCCCTTCAACGCAAGCAGCTGTGTGGTGCACCAT 1019
Db 1055 AAACCTTTCGCAAGCAAAATGGAGCAAGCAATTAATGCCACGGCGGTGTGAGACATAT 1114
QY 1020 GAGGAAGCTACACAT 1034
Db 1115 GAGAAACTACACCT 1129

RESULT 7

AAH25120

ID AAH25120 standard; DNA; 1671 BP.

XX AAH25120;

AC AAH25120;

XX

20-APR-2001; 2001WO-US12992.
20-APR-2000; 2000US-199021P.
28-APR-2000; 2000US-200226P.
05-MAY-2000; 2000US-202339P.
11-MAY-2000; 2000US-203505P.
18-MAY-2000; 2000US-205564P.
26-MAY-2000; 2000US-207739P.
01-JUN-2000; 2000US-208795P.
(INCY-) INCYTE GENOMICS INC.
Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB, Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Wallia NK, Yao MG, Patterson C, Burdill JD, Marcus CA, Zingler KA, Recipon SA, Lu Y, Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR, Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL, Gururajan R;
WPI; 2001-611740/70.
P-PSDB; AAE11768.
Human kinases and nucleic acids, useful for preventing diagnosing and treating cancers, inflammation and immune disorders -
Claim 5; Page 151-152; 166pp; English.
The present invention relates to human kinases (PKIN) and the nucleic acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is used in the prevention, diagnosis and treatment of diseases cancers, adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease, gout, microbial infections, cardiovascular disease and/or inflammation, myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial infarction, cataract, growth and development disorder, seizure disorder, pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage disease, Pick's disease, Tay-Sachs disease, renal disease and obesity. PKIN may be used to treat disorders associated with decreased PKIN expression by rectifying mutations or deletions in a patient's genome that affect the activity of PKIN by expressing inactive proteins or to complement the patients own production of PKIN. PKIN nucleic acids may be used to produce the PKIN polypeptide, by inserting the nucleic acids into a host cell and culturing the cell to express the protein. PKIN nucleic acid and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples and therefore which patients may be in need of restorative therapy. The present sequence is human PKIN-2 cDNA.
Sequence 1736 BP; 463 A; 456 C; 466 G; 351 T; 0 other;
Query Match 20.6%; Score 503; DB 22; Length 1736;
Best Local Similarity 71.2%; Pred. No. 1.3e-131;
Matches 694; Conservative 0; Mismatches 275; Indels 6; Gaps 2;
Qy 66 GGCAATGGTGGGAAGGAAGATGACTGCAGTTCTCTGGAGAAACAGACCACCAACAT 125
Db 155 GGCCATGGCCGGGAGAACGGCAGAGACAGCTCTCTCTGGAAAGACAGCTGAACAT 214
Qy 126 CCGGAAAACCTTCATTTTATGGAAGTGTGGATCAGGAGCTTCTCAGAGCTTTCTCT 185
Db 215 CAAGAAGATCTTCGAGTCTAAAGAGAGCCCTCGGAACGGGGCCCTTTTCCGAGTGGTTT 274
Qy 186 GGTGAAGCAAGAGTACTGGAGGCTTTTGTCTGTAAGTGATC- --AAGAAGTCACC 242
Db 275 AGCTGAAGAGAGCAACTGGCAGCTCTTTGCTGTGAAGTGATCCTTAAGAAGGCGCT 334
Qy 243 TGCCTTCGGGACAGCAGCCTGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGA 302
Db 335 GAAGGGCAAGAAAGCAGCATAGAGATGAGATAGCGCTCTCGAAGAGATTAAAGCATGA 394
Qy 303 AAACATGTGACCTGGAGACATCTATGAGACACACACCACCTACTACTGTGTCATGCA 362
Db 395 AAATATTGTTGCCCTGGGAAGACATTTATGAAGAGCCCAAAATCACTGTACTTGTGTCATGA 454

Qy 363 GCTTGTCTTGTGGGAGACTCTTTGACCGGATCCTGGAGCGGGTGTCTACACAGAA 422
Db 455 GCTGTGTCCGTGGAGAGCTGTTTGACCGGATAGTGGAGAGGGGTTTATACAGAA 514
Qy 423 GGATGCCAGTCTGTGTGATCCAGCAGGTCTTGTGCGCAGTGAATACCTACATGAAATGG 482
Db 515 GGATGCCAGCACTCTGTGATCCCGCAAGTCTTGGACGCCCTCTACTATCTCCACAGATGG 574
Qy 483 CATCTCCACAGAGACTTAAAGCCCGGAAACCTGCTTTACCTTACCCTTGAAGAGACTC 542
Db 575 CATCTCCACAGAGACTTAAAGCCCGGAAATCTCTTGTACTACAGTCAAGATGAGAGTCC 634
Qy 543 TAAGATCATGATCACTGACTTGTGCTGCAAGATGCAACAGAGATG- --CATCATGTC 599
Db 635 CAAATATATGATCACTGACTTGTGATGTTGTCAAAATAGAGGCAAGAGATGTGATGTC 694
Qy 600 CACTGCCGTGGGACCCAGAGCTAGCTGGCTCCAGAACTGTGGCCCAAGAACCTACAG 659
Db 695 CACTGCCGTGGAACTCCAGGCTATGCTGCTCTGAAGTCTCGCCCAAGAACCTTACAG 754
Qy 660 CAAGCTCTGGATTCTGCTCCATCGGCTCATCACCTACATATTGCTCTGTGGATACCC 719
Db 755 CAAAGCCGTGACTGCTGCTCGGATCGGAGTATTGCTCTTGTCTCTCGGCTACCC 814
Qy 720 CCGTCTTATGAAGAACGGAGTCTTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGA 779
Db 815 TCCTTTTATGATGATAAATGACTCCCAAGCTCTTTGAGCAGATCCTCAAGGCGGAATGA 874
Qy 780 GTTTCAGTCTCCATCTCTGGGATGACATTTCTGAGTCAGCCAGGACTTTATTTCACCTT 839
Db 875 GTTTCAGTCTCTCTACTGGGATGACATCTCGGACTCTGCAAAAGACTTTCATTTCGGAACCT 934
Qy 840 GCTTGAGAAGGATCCGAAACGAGCGGTACACCTGTGAGAAGGCGCTTGTAGTCATCCTCGAT 899
Db 935 GATGAGAAGGACCGCGAATAAAGATACACGCTGTGAGCAGCAGCTCGGCCACCATGGAT 994
Qy 900 TCACGGAACAGCGGCCCTCCACCGGAGATCTACCCATGATCAGCTCAGCTCCAGATCCAGAA 959
Db 995 CGCTGGTGACACAGCGCCCTCAACAAAACATCCAGAGTCCGTCAGCGCCAGATCCCGAA 1054
Qy 960 GAACTTTGCTAAGCAAGTGGAGGCAAGCCTTCAACGCAAGCAGCTGTGGTGCACCAT 1019
Db 1055 AAACCTTCCCAAGACCAATGGAGACAGCATTTAATCCACGCGCGCTCGTGAGACATAT 1114
Qy 1020 GAGGAAGTACACAT 1034
Db 1115 GAGAAAACCTACACCT 1129
RESULT 10
AAI60424
ID AAI60424 standard; cDNA; 2164 BP.
XX AC AAI60424;
XX AC AAI60424;
XX 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 4413.
XX Human; nucleotopic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia; ss.
XX Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.

PF	26-DEC-2000; 2000WO-US34263.		
XX	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX	(HYSE-) HYSEQ INC.		
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QA, Zhou P, Goodrich R, Dimanac RT;		
XX	WPI; 2001-442253/47.		
DR	P-PSDB; AAM41268.		
XX	Novel nucleic acids and polypeptides, useful for treating disorders		
PT	such as central nervous system injuries -		
XX	Claim 1; SEQ ID NO 4413; 10078pp; English.		
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and		
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities such as: Immune system suppression,		
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic		
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S disorders.		
CC	Note: The sequence data for this patent did not form part of the printed		
XX	specification.		
XX	Sequence 2164 BP; 557 A; 542 C; 560 G; 505 T; 0 other;		
SQ	Query Match 20.6%; Score 503; DB 22; Length 2164;		
	Best Local Similarity 71.2%; Pred. No. 1.4e-131;		
	Matches 694; Conservative 0; Mismatches 275; Indels 6; Gaps 2;		
QY	66 GGCAATGGGTGCGAAGAGAGATGACTGCGAGTTCCTCGAAGAAACAGACCACCAACAT 125		
DB	140 GGCCATGGCCCGGAGAGCGCGAGAGAGCTCTCTCGAAGAAAGCAAGCTGAAGACAT 199		
QY	126 CCGGAATCTTCAATTTTATGGAAGTGTGGAGTCAGGAGCTTCTCAGAAAGTTTCT 185		
DB	200 CAAGAAGATCTTCGAGTTCAGAGAGACCTCGGAACCGGGGCTTTTCGGAAGTGGTTT 259		
QY	186 GGTGAACGAAGACTGACTCGGAGCTCTTCTCTGAGAGTGCATC---AAGAAGTCACC 242		
DB	260 AGCTGAAGAGAGCGGCACTGGCAGCTCTTTGCTGTAAGATGTATCCCTAAGAAGCGGT 319		
QY	243 TGCCTTCGGGACAGACGCTGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCATGA 302		
DB	320 GAAGGCAAGGAAGCAGCATAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGA 379		
QY	303 AACATTGTGACCTCGAGAGACATCTATGAGAGACACACCCACTACTACTGTGTATGCA 362		
DB	380 AAATATTGTTCCTCGAAGACATTTATGAAAGCCCAATACACCTGTACTTGTGTATGCA 439		
QY	363 GCTGTGTTCTGTTGGGAGCTCTTTGACCGGATCCTTGGAGCGGGTGTCTACACAGAA 422		
DB	440 GCTGGTTCGGGTGGAGAGCTGTTTACCAGGATGTTGAGAGAGGGGTTTATACAGAGAA 499		
QY	423 GGATGCCAGTCTGTTGATCCAGCAGGTCTTGTGCGCAGTGAATACCTACATGAGAAATGG 482		
Db	500 GGATGCCAGCACTCTGATCCGCCAAGTCTTGGAGCGCGTGTACTATCTCCACAGAATGGG 559		
QY	483 CATCGTCCACAGAGACTTTAAAGCCCGAAACCTGCTGCTTTACCTTACCCCTGAAGAGAACTC 542		
Db	560 CATCGTCCACAGAGACCTCAAGCCCGAAATCTCTTGTACTACAGTCAAGATGAGGAGTC 619		
QY	543 TAAGATCATGATCACTGACTTTGGTCTGTCAAGATGGAACAGAATGG---CATCATGTC 599		
Db	620 CAAATATGATGATGACTGCTTGGATGTCTAAATATGAGGGCAAGAGATGTGATGTC 679		
QY	600 CACTGCTGTGGGACCCAGGCTAGTGGTCTCAGAAAGTGTGCGCCAGAAACCTTACAG 659		
Db	680 CACTGCTGTGGAACTCCAGGCTATGCTGCTCTGAAAGTCTCTGCGCCAGAAACCTTACAG 739		
QY	660 CAAGGCTGTGGATTGCTGCTGCTCAGGCTCATCACATATATTGCTCTGTGGATACCC 719		
Db	740 CAAAGCGGTTGACTGCTGGTCCATCGGAGTGAATGCTTGTGCTCTGCGGCTACCC 799		
QY	720 CCCGTTCTATCAAGAAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGGCTACTATGA 779		
Db	800 TCCITTTTATGATGAAATGACTCCAAAGCTCTTTGAGCAGATCTCAAGGCGGAATATGA 859		
QY	780 GTTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAAGCAAGGACTTTATTTGCCACTT 839		
Db	860 GTTTGACTCTCCCTACTGGGATGACATCTCGGACTCTGCAAAAGACTTCATTCGGAACCT 919		
QY	840 GCTTGAGAAGATCCGACAGCGGTCACCTGTGGAAGAGCCCTTGAGTCAATCCCTGGAT 899		
Db	920 GATGAGAAGAGCCGCAATAAAGATACACGTGTGAGCAGGAGCTCGGACCCCATGGAT 979		
QY	900 TGACGGAAACACGCGCTCCACCGGGACATCTACCCATCAGTCAAGCTCCAGATCCAGAA 959		
Db	980 CGCTGGTGACACAGCCCTCAACAAAACATCCAGAGTCCGTCAGCCGCCAGATCCGGAA 1039		
QY	960 GAACCTTTGCTAAGAGCAAGTGGAGGCAAGCCCTTCAAGCAGCAGCAGCTGTGGTGACACCAT 1019		
Db	1040 AAACCTTTGCCAAGAGCAAAATGGAGCAAGCATTTAATGCCAGCGCGCTGCTGACACATAT 1099		
QY	1020 GAGGAAGCTACACAT 1034		
Db	1100 GAGAAACTACACCT 1114		
RESULT 11			
AAS06708			
ID	AAS06708 standard; cDNA; 1074 BP.		
XX	AAS06708;		
XX	12-SEP-2001 (first entry)		
XX	Polynucleotide sequence encoding human protein kinase #8.		
DE	Human; protein kinase; PTK; STK; cancer; cardiovascular disease;		
XX	metabolic disorder; immune related disease; neurological disorder;		
KW	neurodegenerative disorder; inflammatory disorder; infectious disease;		
KW	reproductive disorder; gene therapy; ss.		
XX	Homo sapiens.		
OS	WO200138503-A2.		
PN	31-MAY-2001.		
PD	22-NOV-2000; 2000WO-US32085.		
XX	24-NOV-1999; 99US-0167482.		
XX	(SUG-) SUGEN INC.		
PA	Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;		
PI	Flanagan P, Clary D;		

XX WPI: 2001-343950/36.
DR P-PSDB; AAU03508.
XX
XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX
XX Example 1; Figure 1; 433pp; English.
XX
CC RAS06701-AAS06757 encode for novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.
XX
XX Sequence 1074 BP; 313 A; 258 C; 276 G; 227 T; 0 other;
SQ

Query Match 20.5%; Score 502.2; DB 22; Length 1074;
Best Local Similarity 71.3%; Pred. No. 1.7e-131;
Matches 692; Conservative 0; Mismatches 273; Indels 6; Gaps 2

QY 70 ATGGGTCGAAGGAAGACATGACTGCTTCCTGGAGAAGACAGACCACCATCCGG 129
DB 1 ATGGCCCGGGAGACGGGAGAGACGCTTCCTCGTGAAGAAAGCTGAAGACATCAAG 60
QY 130 AAACACTTCATTTTATGGAAGTGTGGGATCAGAGCTTCTCAGAAAGTTTCTCGTGGT 189
DB 61 AAGACTCTCCAGTTCAAAGAGACCTCGGAACCGGGGCTTTTCCGAAGTGGTTTAACT 120
QY 190 AAGCAAGACTGACTGGGAGCTCTTCTCTGAGTGCATC ---AAGAAGTCACCTGCC 246
DB 121 GAAGAGAGGCAACTGGCAAGCTTTGCTGTGAAGTGATATCCCTAGAAGGGCGCTGAAG 180
QY 247 TTCGGGGACAGACGCTCGAGAAATCAGATGCTGTGTGTTGAAAAAGATCAAGCATGAAAC 306
DB 181 GGCAAGGAAGCAGCATAGAGAATCAGATAGCCGCTCTGAGAAAGATTAAGCATGAAAT 240
QY 307 ATTGTGACCTTGAGGACATCTATGAGAGACACCCCACTACTACTTGTGTCATGCAGCTT 366
DB 241 ATTGTTGCCCTGGAAGACATTTATGAAAGCCCAATCACTCTACTTGTGTCATGCAGCTG 300
QY 367 GTTTCCTGGTGGGAGCTCTTTCACCGCATCTCTGGAGCGGGTGCTACACAGAGAAGGAT 426
DB 301 GTGTCGGTGGAGAGCTGTTTGCCCGGATGATGGAGAGGGGTTTTATACAGAGAAGGAT 360
QY 427 GCCAGTCTGGTGATCCAGCAGGCTCTTGTGGCAGTGAATACCTACATGAGAATGGCATC 486
DB 361 GCCAGCACTCTGATCCGCCAAGCTTTGGACGCGGTGTACTATCTCCACAGATGGGCATC 420
QY 487 GTCCACAGAGACTTAAGCCCGAAAACCTGCTTTACCTTACCTTACCCTGGAAGAACTCTAAG 546
DB 421 GTCCACAGAGACTTCAAGCCCGAAAATCTCTGTGTAAGTCAAGATGAGAGATGCCAAA 480
QY 547 ATCATGATCACTGACTTTTGGTGTCTGCAAGATGGAACAGAATGG ---CATCATCTCCACT 603
DB 481 ATAATGATCAGTGACTTTGGATGTGCAAAATGAGGCAAGAGAGATGTGATGTCCTACT 540
QY 604 GCCTGTGGGACCCCGAGCTTACGTGGCTCCAGAAAGTGTGGCCCAAGAACCCCTACAGCAAG 663
DB 541 GCCTGTGGAATCTCAGGCTATGCTCGCTCCTCAAGTCTTCGCCCAAGAACCTTACGAAA 600

Qy	664	GCTGTGATTTCGTGGTCCATCGGGCGTCATCACCTACATATTGCTCTGTGGATACC	CCCCCG	72
Dd	601	GCCTTGACTGCTGGTCCATCGGAGTGATTCCTACATCTTGCTCTGCGCTAC	CCTCCT	66
Qy	724	TTCCTATCAAGAACCGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGTTT	783	
Dd	661	TTTTATGATGAATAATGACTCCAAGCTCTTTGAGCAGATCCTCAAGCGGGAATATGAGTTT	720	
Qy	784	GAGTCTCCATCTCTGGGATGACATTTCTTGAGTCAAGCAAGGACTTTTATTTGCCACTTGCTT	843	
Dd	721	GACTCTCCCTACTGGGATGACATCTCGACTCTTGCAAAAGACTTCATTTCGGAACCTGATG	780	
Qy	844	GAGAAGGATCCGAACGAGCGGTACACTGTGGAAGGCGCTTGAGTCAATCCCTGGATGAC	903	
Dd	781	GAGAAGGACCGCAATAAAGATACAGCTGTGAGCAGCGCTCGGCACCATGGATCGCT	840	
Qy	904	GGAAACACGCGCTCCACCGGGACATCTACCCATCAGTCAGCTCCAGATCCAGAAAGAAC	963	
Dd	841	GGTGACACAGCCCTCAACAANAACATCCAGAGTCCGTGAGCCCGAGATCCGGAANAAC	900	
Qy	964	TTTGCTAAGAGCAAGTAGGGCAAGCGCTTCAACGCGACGAGCTGTGTGTGCCACCATGAGG	1021	
Dd	901	TTTGCCAAAGAGCAAATGAGACAAAGCATTTAATGCCACGCGCTCGTGAGACATATGAGA	960	
Qy	1024	AAGCTACACAT	1034	
Dd	961	AAACTACACCT	971	
 RESULT 12 AAH25119				
ID	AAH25119 standard; DNA; 1074 BP.			
XX	AAH25119;			
XX	22-AUG-2001 (first entry)			
DE	Nucleotide sequence of a human kinase polypeptide.			
KW	Human; kinase; human disease; human disorder; ss.			
OS	Homo sapiens.			
FH	Key Location/Qualifiers			
FT	CDS 1..1074			
FT	/*tag= a			
FT	/transl_except= "(pos: 838..843, aa: Ala)"			
XX	/product= "Kinase"			
PN	WO200142435-A2.			
XX	14-JUN-2001.			
PD				
XX	07-DEC-2000; 2000WO-US33240.			
Pf				
XX	07-DEC-1999; 99US-0169428.			
PR				
XX	(LEXI-) LEXICON GENETICS INC.			
PA				
XX	Donoho G, Scoville J, Turner CA, Friedrich G, Zambrowicz B;			
PI	Abuin A, Sands AT;			
FI				
XX	WPI; 2001-381667/40.			
DR	P-PSDB; AAB84360.			
DR				
XX	Novel isolated human kinase polynucleotide that shares structural			
PT	similarity with animal kinases including calcium/calmodulin-dependent			
PT	protein kinases and serine/threonine protein kinases, useful in			
PT	therapeutics -			
XX	Disclosure; Page 30-31; 32pp; English.			
PS				
XX				

XX	DT	04-JUL-2001 (first entry)	
XX	DE	Human death domain-containing receptor (DDCR) cDNA from HDPV291 clone.	
XX	KW	Human; death domain-containing receptor; DDCR; immunosuppressive;	
KW	antiarthritic; antirheumatic; antiproliferative; cytostatic;		
KW	cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;		
KW	antibacterial; virucide; fungicide; ophthalmological; gene therapy;		
KW	immunodeficiency disease; Acquired immune deficiency syndrome;		
KW	AIDS; leukaemia; autoimmune disease; systemic lupus erythematosus;		
KW	hyperproliferative disorder; neoplasm; cerebrovascular disorder;		
KW	cerebral ischaemia; angiogenesis; cardiovascular disorder;		
KW	neurodegenerative disease; Alzheimer's disease; Parkinson's disease;		
KW	ocular disorder; corneal infection; degenerative disease; SMA; apoptosis;		
KW	spinal muscular atrophy; epithelial cell proliferation; infection;		
KW	cancer; wound healing; skin aging; chemotaxis; HDPV291 clone; ss.		
XX	XX	Homo sapiens.	
OS	XX		
XX	Key	Location/Qualifiers	
FT	CDS	288..1130	
FT		/*tag= a	
FT		/product= "Human death domain-containing receptor"	
FT			
XX	XX	WO200129063-A2.	
XX	PD	26-APR-2001.	
XX	XX	17-OCT-2000; 2000WO-US28666.	
XX	XX		
PR	18-OCT-1999;	99US-0159585.	
PR	24-NOV-1999;	99US-0167246.	
XX	XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	XX		
XX	XX	NI J, Ruben SM;	
PI	XX	WPI: 2001-300316/31.	
DR	DR	P-PSDB; AAE01034.	
XX	XX	New death-domain containing receptor polynucleotides and polypeptides,	
PT	PT	useful for treating and diagnosing cancer	
PS	XX	Claim 1; Page 283-284; 298pp; English.	
XX	XX	The patent discloses novel death domain-containing receptor (DDCR)	
CC	CC	cDNAs and their corresponding proteins. DDCR cDNA and protein are	
CC	CC	used to prevent, treat or ameliorate a medical condition in mammals.	
CC	CC	They are also used in diagnosing a pathological condition or	
CC	CC	susceptibility to a pathological condition. The DDCR protein and	
CC	CC	its antibodies are used in the diagnosis and treatment of disorders	
CC	CC	such as immunodeficiency diseases (e.g. Acquired immune deficiency	
CC	CC	syndrome (AIDS), leukaemia) autoimmune diseases (e.g. systemic lupus	
CC	CC	erythematosus, rheumatoid arthritis), hyperproliferative disorders	
CC	CC	(e.g. neoplasms of the breast or liver), cerebrovascular disorders	
CC	CC	(e.g. cerebral ischaemia, angiogenesis), cardiovascular disorders	
CC	CC	(e.g. cardiac arrest), neurodegenerative diseases (e.g. Alzheimer's	
CC	CC	disease, Parkinson's disease), ocular disorders (e.g. corneal	
CC	CC	infection), degenerative diseases (e.g. spinal muscular atrophy-SMA),	
CC	CC	cancer, aberrant apoptosis, disorders of the placenta or uterus and	
CC	CC	infections caused by bacteria, viruses and fungi. The DDCR proteins	
CC	CC	are used to aid wound healing and epithelial cell proliferation,	
CC	CC	to prevent skin aging due to sunburn, to maintain organs before	
CC	CC	transplantation, to support cell culture of primary tissues, to	
CC	CC	regenerate tissues and in chemotaxis. The DDCR sequences are used	
CC	CC	in correcting aberrant cellular apoptosis by gene therapy.	
CC	CC	The present sequence is human death domain-containing receptor	
CC	CC	(DDCR) cDNA from clone HDPV291 (ATCC No: PTA839).	
XX	XX		
SQ		Sequence 1578 BP; 432 A; 381 C; 401 G; 364 T; 0 other;	
		Query Match 20.3%; Score 497; DB 22; Length 1578;	

Query Match 20.3% Score 497; DB 22; Length 1578;

Best Local Similarity 71.6%; Pred. No. 6.1e-130;		Matches 682; Conservative 0; Mismatches 265; Indels 6; Gaps 2;	
QY	88	GATGACCTGCTGCTCGAAGAAACACACACCAACATCCGGAAACCTTCATTTTATG	147
Db	15	GAGAGCAGCTCCCTCGTGAAGAAACCAAGCTGAAGACATCAAGAGATCTTCGAGTTCAA	74
QY	148	GAAGTGTGGGATCAGAGCTTTCTCAGAAGTTTCTCGTGAAGCAAGACTGACTGGG	207
Db	75	GAGACCTCGAAGCGGGCTTTTCGAAGTGGTTTAGCTCAAGAGAGGCAACTGGC	134
QY	208	AAGCTCTTTGCTGTAAGTGCATC---AAGAAGTCACTGCTCCGCGACAGACGCTG	264
Db	135	AAGCTCTTTGCTGTAAGTGCATC---AAGAAGTCACTGCTCCGCGACAGACGAT	194
QY	265	GAGAAATGAGATGCTGTTGAAAGAGATCAAGCATGAAGAAATTTGACCCCTGGAGGAC	324
Db	195	GAGAAATGAGATGCTGTTGAAAGAGATTAAGCAATGAAGAAATTTGACCCCTGGAGGAC	254
QY	325	ATCTATGAGAGCACACCCACTACTACCTGGTGCATGAGCTTTCTTGGTGGGAGCTC	384
Db	255	ATTTATGAAGCCCAATACCTGTACTTGGTGCATGAGCTGGTGGTGGAGAGCTG	314
QY	385	TTTGACCGGATCTGGAGCGGGTGCTACACAGAGAAAGATGCCAGTCTGTGTGATCCAG	444
Db	315	TTTGACCGGATCTGGAGCGGGTGCTTATACAGAGAAAGATGCCAGACTCTGTATCCCG	374
QY	445	CAGTCTTTGCGCAGTGAATACCTACATGAGATGGATCGTCCACAGAGACTTAAAG	504
Db	375	CAAGTCTTTGCGCAGTGAATACCTACATGAGATGGATCGTCCACAGAGACTTAAAG	434
QY	505	CCGAAACCTGCTTTACCTTACCCCTGAAGAGAACTCAAGATCATGATCACTGACTTT	564
Db	435	CCGAAATCTCTGTACTACATCAAGATGAGGATCCAAATATATGATCACTGACTTT	494
QY	565	GGTCTGTCGAAGTGAACAGATGG---CATCATGTCCACTGCCCTGGGACCCGAGG	621
Db	495	GGATTGTCAAAATGGAGGCAAGGAGATGTGATGTCCACTGCCCTGGGACCCGAGG	554
QY	622	TACGTGCTCCAGAGTGTGGCCAGAAACCTCAGCAAGGCTGTGGATTGCTGCTCC	681
Db	555	TATGTCCTCTGAGTCTCGCCAGAAACCTTACAGAAACCGGTGACTGCTGCTGCTC	614
QY	682	ATCGGCTCATACCTACATATTTGCTGTGGATACCCCGCTTCTATGAAGAAACGAG	741
Db	615	ATCGGAGTATGCTTACATCTTGTCTGCGGCTACCCCTCTTTTATGATGAATGAC	674
QY	742	TCTAAGCTTTTCGAGAGATCAAGGAGGCTACTATGAGTTTGAGTCTCCATTTCTGGAT	801
Db	675	TCCAAGCTCTTTGAGCAGATTCCTCAAGCGGGAATATGAGTTTGACTCTCCCTACTGGAT	734
QY	802	GACATTTCTGAGTCAGCAAGGACTTTATTTGCCACTTGCTTCAAGAGGATCCGAACGAG	861
Db	735	GACATCTCCGACTCTGCAAGACTTCATTCGGAACCTGATGGAGAGGACCCGATATA	794
QY	862	CGGTACACCTGTGAGAGGCTTTGAGTCAATCCCTGGAATGACGGAACACGCGCTCCAC	921
Db	795	AGATACACGTGTGAGCAGGAGCTCGGCACCCCATGGATCGTGGTGCACAGCCCTCAAC	854
QY	922	CGGGACATCTACCCATCAGTCAAGCTCCAGATCCAGAGAACTTTGCTTAAGAGCAAGTGG	981
Db	855	AAAACATCACGAGTCCGTCAGCGCCAGATCCGGAACAACTTTGCCAAGAGCAATGG	914
QY	982	AGGCAAGCTTTCAACGAGCAGCTGTGGTGCACCATGAGGAGCTACACAT	1034
Db	915	AGCAACCATTTTATGCCACGGCGCTGCTGAGACATATGAGAAACTACACCT	967

RESULT 15
AAC90432
ID AAC90432 standard; cDNA; 1694 BP.
XX
AC AAC90432;

XX	19-MAR-2001 (first entry)	
DT	Murine Dendritic Cell Kinase 3 coding sequence.	
DE	Murine Dendritic Cell Kinase; MDCK-3; autoimmune disorder; allergy;	
XX	Murine; Dendritic Cell Kinase; MDCK-3; autoimmune disorder; allergy;	
KW	wound healing; periodontal disease; inflammatory disease; tumour;	
KW	infection; Ca2+/calmodulin-dependent kinase family; ss.	
XX	Mus musculus.	
OS	WO200073468-A1.	
XX	07-DEC-2000.	
PN	26-MAY-2000; 2000WO-US14696.	
XX	28-MAY-1999; 99US-0136781.	
PD	(IMMV) IMMUNEX CORP.	
XX	Bird TA, Virca GD, Martin U, Anderson DM;	
PI	WPI; 2001-061546/07.	
DR	P-PSDB; AAB50055.	
XX	Novel murine and human kinase nucleic acids useful for treating	
PT	inflammations, infections, tumors, allergies, autoimmune diseases, and	
PT	for stimulating or suppressing Immune responses -	
XX	Claim 1; Page 86; 106pp; English.	
PS	The present sequence is the coding sequence for Murine Dendritic Cell	
XX	Kinase 3 (MDCK-3). The protein encoded by the present sequence is useful	
CC	for treating a variety of disorders listed in the disclosure of the	
CC	specification, including autoimmune disorders, allergic reactions,	
CC	myeloid or lymphoid cell deficiencies, wound healing and tissue repair	
CC	and replacement, burns, incisions and ulcers, periodontal disease,	
CC	inflammatory diseases, tumours and bacterial, viral or fungal infection.	
CC	MDCK-3 is a member of the Ca2+/calmodulin-dependent kinase family.	
XX	Sequence 1694 BP; 437 A; 444 C; 464 G; 349 T; 0 other;	
SQ	Query Match 20.1%; Score 492.6; DB 22; Length 1694;	
	Best Local Similarity 70.6%; Pred. No. 1.1e-128;	
	Matches 686; Conservative 0; Mismatches 279; Indels 6; Gaps 2;	
QY	67	GCAATGGTGCAGGAAGAGATGACTGCTGAGTTCCTGGAAGAAACAGCACCACATC 126
Db	240	GCCATGGCCGGGAGAACGGCGGAGAGAGCTCCTCTGGAAACCAAGCAGAGACATT 299
QY	127	CGGAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAAGTTTCTCTG 186
Db	300	AAGAAGATCTTCGAGTTCAAGGAGACCCCTCGAACTGGGGCTTTCTCGAAGTTGTTTA 359
QY	187	GTGAAGCAAAAGACTGACTGGGAAGCTCTTCTGCTGAAGTGCATC---AAGAAGTCACT 243
Db	360	GCCGAGGAGAAAGCTACTGGGAAGCTCTCCGAGTGAAGTGCATCCCGAAGAGGGCTG 419
QY	244	GCTTCCGGGACAGCAGCCTGGAGAAATGAGATTGCTGTGTTGAAAGAAATCAAGCATGA 303
Db	420	AAGGCGAAGGAGAGCAGCATCGAGAACGAGATTGCCGTGCTTAGAAAGATTAAAGCATGA 479
QY	304	AACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCACTACTACCTGGTTCATGCAG 363
Db	480	AACATTGTGCTTGGAGATATTTATGAAAGCCCAATACCTCTACCTCTGCTCATGCAA 539
QY	364	CTTGTCTTCTGTTGGGAGCTCTTTGACCGGATCTTGGAGCGGGGTGTCTACACAGAGAAG 423
Db	540	CTTGTGCTGTTGGAGAACTCTTCGATCGGATAGTGGAGAGGGGTTTTACACAGAGAA 599
QY	424	GATGCCAGTCTGTTGATCCAGCAGCTCTTGTGCGGAGTGAATACCTACATGAGAATGCC 483

Db 600 GATGCCAGACCTCTCATCGCCAGGCTCTGGATGCCGTATACTATCTCCACAGAAATGGGC 659
QY 484 ATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTCAAGAGAACTCT 543
Db 660 ATTGTCACAGGGACCTCAAGCCGAGAACTCTTTATACAGTCAAGACGAGAGTCC 719
QY 544 AAGATCATGATCAGTACTGTTGGTCTGTCCAAAGATGGAACAGAAATGG--CATCATGTCC 600
Db 720 AAAATAATGATCAGTACTGTTGGTCTGTCCAAATGGAGGGCAAGAGAGATGTGATGTC 779
QY 601 ACTGCTGTGGGACCCAGGCTACCTGGCTCCAGAGTGCCTGGCCACAAAACCTTACAGC 660
Db 780 ACGGCTCGGGGACCCAGGCTATGTGCTCGGAAAGTCTCGCCCCAGAAACCCGTACAGC 839
QY 661 AAGGGTGTGGATGCTGTCCATCGCGCTCATCACCTACATATGCTCTGTGGATACCCC 720
Db 840 AAAGCTGTGACTGTGTTCCATCGGGGTGATCGCTATATCTGCTCTGTGGTTACCCCT 899
QY 721 CCGTTCATGAGAAGCGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGGCTACTATGAG 780
Db 900 CCTTTTATGATGAAAATGACTCGAAGCTGTTTGAACAGATCCTCAAGGCAGAAATATGAG 959
QY 781 TTTGAGTCTCCATTCTGGGATGACATTTCTGAGTCAGCCAAAGGACTTTATTTGCCACTTG 840
Db 960 TTTGATTCCTCTACTGGGATGACATCTCCGACTCTGCCAAAGACTTCATTTCGGAATCTG 1019
QY 841 CTTGAGAAGGATCCGAACGAGCGGTACACCTGTGAGAAGGCTTGTGAGTCATCCCTGGATT 900
Db 1020 ATGGAGAAGACCCCAATAAAGATACACTTGTGAGCAGGAGCTCGACACCCATGGATT 1079
QY 901 GACGGAACACGGCCCTCCACCGGACATCTACCCAGTCAGCTCCAGATCCAGAAG 960
Db 1080 GCTGTTGACACAGCCCTTAGCAAAACATTCACGAATCTGTCACTGCCAGATCCGGAAG 1139
QY 961 AACTTTGCTAAGCAAGTGGAGGCAAGCTTCAACGAGCAGCTGTGTCACACATG 1020
Db 1140 AATTTTGAAGAGCAAAATGGAGACAAGCGTTTAAACGCCAGGCAGTGTGAGACATATG 1199
QY 1021 AGGAAGCTACA 1031
Db 1200 CGGAGGCTCCA 1210

Search completed: March 14, 2003, 14:17:14
Job time : 574 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 11:30:44 ; Search time 3881 Seconds
(without alignments)
10211.378 Million cell updates/sec

Title: US-09-960-643-1
Perfect score: 2447
Sequence: 1 tggatggagctcaagcag.....ttttctctaaaaa 2447

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	Length	DB ID	Description
1	963.8	39.4	1126	13	BM547443	BM547443 AGENCOURT
2	826	33.8	966	13	BI084897	BI084897 602869466
3	801.4	32.8	1048	13	BI084101	BI084101 602869466
4	794.2	32.5	809	13	BI821474	BI821474 603038366
5	736.8	30.1	740	13	BI824483	BI824483 603038855
6	720.8	29.5	742	13	BI818261	BI818261 603032510

7	720.4	29.4	1068	14	BM921532	BM921532 AGENCOURT
8	713.2	29.1	812	13	BI772626	BI772626 603060879
9	710.2	29.0	1129	14	BM807335	BM807335 AGENCOURT
10	701.4	28.7	824	12	BG715920	BG715920 602676667
11	699	28.6	731	13	BI753035	BI753035 603025844
12	653.6	26.7	905	14	BQ934044	BQ934044 AGENCOURT
13	590.6	24.1	818	12	BG293660	BG293660 602390529
14	555	22.7	656	13	BI667965	BI667965 603292877
15	522	21.3	554	13	BI834635	BI834635 603090418
16	506.8	20.7	573	14	BQ086330	BQ086330 1121607.y
17	499.4	20.4	501	9	AL134342	AL134342 DKFZp5470
c	479	19.6	479	14	BQ102407	BQ102407 1121607.x
19	476.2	19.5	1051	9	AL560091	AL560091 AL560091
c	473.2	19.3	484	10	AW016039	AW016039 UI-H-BI0p
c	454.2	18.6	555	10	AW251224	AW251224 UI-R-BJ0-
22	441.4	18.0	620	14	BM944418	BM944418 UI-M-EH0p
c	410	16.8	410	9	AA838372	AA838372 of29a11.s
24	405.8	16.6	962	9	AL556476	AL556476 AL556476
25	397.2	16.2	918	14	BQ949629	BQ949629 AGENCOURT
26	378.2	15.5	873	12	BG186668	BG186668 602319906
c	371	15.2	451	10	AW254051	AW254051 UI-R-BJ0-
28	360.4	14.7	929	13	BI758466	BI758466 603022650
c	358.6	14.7	439	10	AW522224	AW522224 UI-R-BJ0-
30	353	14.4	871	10	BE569540	BE569540 601328368
31	347.4	14.2	1068	14	BM919261	BM919261 AGENCOURT
32	346.2	14.1	627	10	AW826802	AW826802 FK53602.y
33	345.8	14.1	493	10	AW520558	AW520558 UI-R-BJ0p
34	330.8	13.5	798	12	BG702279	BG702279 602683496
c	328.8	13.4	941	9	AL578723	AL578723 AL578723
36	327	13.4	822	13	BI545545	BI545545 603187723
c	326.8	13.4	978	9	AL574819	AL574819 AL574819
38	325.2	13.3	556	12	BG895406	BG895406 358836 MA
39	324.6	13.3	1052	14	BQ066920	BQ066920 AGENCOURT
40	321.8	13.2	659	13	BI463145	BI463145 603207913
41	321.8	13.2	758	13	BI761267	BI761267 603044016
42	321.6	13.1	636	10	BE266955	BE266955 601191504
43	321.2	13.1	645	14	BM728430	BM728430 UI-E-EJ0-
44	319.2	13.0	580	9	AJ397261	AJ397261 AJ397261
45	318.8	13.0	748	12	BF143967	BF143967 601791217

ALIGNMENTS

RESULT 1
BM547443
LOCUS
DEFINITION
AGENCOURT_6507688 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5724450
5', mRNA sequence.
ACCESSION
BM547443
VERSION
BM547443.1 GI:18781213
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1126)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM12713 Row: 1 Column: 19
High quality sequence stop: 623.
Location/Qualifiers
1..1126
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5724450"
/clone_lib="NIH_MGC_125"
/lab_host="DH10B"

/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dr primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

BASE COUNT 299 a 291 c 286 g 244 t 6 others
ORIGIN

Query Match 39.4%; Score 963.8; DB 13; Length 1126;
Best Local Similarity 94.3%; Pred. No. 3.le-246;
Matches 1063; Conservative 0; Mismatches 53; Indels 11; Gaps 6;

QY 2 GGAGTGGGAGCTCAACGACGATTCTTCCCGAGTCCCTGSCATCCTCAGAAGCTTCAACTC 61
DB 1 GGAGTGGGAGCTCAACGAGATTCTTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACTC 60
QY 62 TGGAGGCAATGGTTCGAAAGGAAGATGACTGCAGTTCCTCGAAGAAACAGACCA 121
DB 61 TGGAGGCAATGGTTCGAAAGGAAGATGACTGCAGTTCCTCGAAGAAACAGACCA 120
QY 122 ACATCGGGAACCTTCATTTTATGGAAGTCTGGGATCGGAGCTTTCTCAGAAGTTT 181
DB 121 ACATCGGGAACCTTCATTTTATGGAAGTCTGGGATCGGAGCTTTCTCAGAAGTTT 180
QY 182 TCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTGTCTCTGAAGTGCATCAAGAAGTCAC 241
DB 181 TCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTGTCTCTGAAGTGCATCAAGAAGTCAC 240
QY 242 CTGCTTCGGGACAGACGCTGGGAATGAGATTGCTGTGTGTTGAAAAGATCAAGCATG 301
DB 241 CTGCTTCGGGACAGACGCTGGGAATGAGATTGCTGTGTGTTGAAAAGATCAAGCATG 300
QY 302 ARAACATTGTACCTGGAGGACATCTATGAGACACCAACCTACTACTCTGGTCTATGC 361
DB 301 ARAACATTGTACCTGGAGGACATCTATGAGACACCAACCTACTACTCTGGTCTATGC 360
QY 362 AGCTTTCTGTGTGGGAGCTTTTGACCGGATCTTGAGCGGGGTGTCTACACAGAGA 421
DB 361 AGCTTTCTGTGTGGGAGCTTTTGACCGGATCTTGAGCGGGGTGTCTACACAGAGA 420
QY 422 AGATGCCAGTCTGGTGATCCAGCAGGCTTTCTCGGCGAGTGAATACCTACATGAGAATG 481
DB 421 AGATGCCAGTCTGGTGATCCAGCAGGCTTTCTCGGCGAGTGAATACCTACATGAGAATG 480
QY 482 GCATCTGTCACAGAGACTTAAGCCCGAAGAACTCTTACCTTACCTTACCTTGAAGAGACT 541
DB 481 GCATCTGTCACAGAGACTTAAGCCCGAAGAACTCTTACCTTACCTTGAAGAGAACT 540
QY 542 CTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCCA 601
DB 541 CTAAGATCATGATCACTGACTTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCCA 600
QY 602 CTGCTGTGGGACCCAGGCTACGTGGGTCCAGAGTGTGGCCAGAAACCTTACAGCA 661
DB 601 CTGCTGTGGGACCCAGGCTACGTGGGTCCAGAGTGTGGCCAGAAACCTTACAGCA 660
QY 662 AGGCTGTGATGCTGGTTCATCGCGTCATCACCTACATATGCTCTGTGGATACCC 721
DB 661 AGGCTGTGATGCTGGTTCATCGCGTCATCACCTACATATGCTCTGTGGATACCC 720
QY 722 CGTTCTATGAAGAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGT 781
DB 721 CATTCTATCAAGAACGGAGTCTAAGCTTTTCGAGAAGATCAAGGAGGCTACTATGAGT 780

QY 782 TTGAGTCTCCATTTCTGGGATGACATTTCTTGAGTCAGCCAGGACTTATTGTCAC--TTG 840
DB 781 TTGAGTCTCCATTTCTGGGATGACATTTCTTGAGTCAGCCAGGACTTATTGTCAC--TTG 840
QY 841 CTTGAGGAGTCCGACGAGCGGTACACCTGT--GAGAAGGCCCTTGAGTCATCCC--TG 896
DB 841 CTTGAGGAGTCCGACGAGCGGTACACCTGTGAGAAAGGCCCTTGAGTCATCCCCTGG 900
QY 897 GATTGACGGAACACAGGCGC--TCCACGCGGACATCTACCCATCAGTCAGCCTCCAGATC 954
DB 901 GATTGACGGAACACAGGCGCCTCCACGGGNACATCTACCCATCAGTCAGCCTCCAGATC 960
QY 955 CAGAAGAACTTTGTTAAGAGCAAGTGGAGCAAGCCTTCAACGAGCAGAGCTGTGGTGCAC 1014
DB 961 CAGAAGAACTTTGGCTTAANAACCAATGGAGGACGCTTCAACCCANCAAGCTGTGGTGCC 1020
QY 1015 CACATGAGGAGCTTACATGAACTGCACAGCCCGGCGTCCGCC--AGAGGTGGAG 1071
DB 1021 CCCCTGAGGAAGCTAC--CCTGAACCTGGACAGCGGGAATTCNNCCCCAAAGGGGGA 1079
QY 1072 AACAGGCGCCTGAAACTCAAGCCTCAGAAACCTCTAGACCCAGCTC 1118
DB 1080 AACAGGCGCCTGAATCCAGCCTCAAAACCTTTGAACCCAGTTC 1126

RESULT 2
BI084897/c 966 bp mRNA linear EST 20-JUN-2001
LOCUS 602869466T1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013856 3',
DEFINITION mRNA sequence.
ACCESSION BI084897
VERSION BI084897.1 GI:14503227
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 966)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1820 row: d column: 17
High quality sequence stop: 832.

FEATURES
source Location/Qualifiers
1..966
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5013856"
/clone_lib="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dr priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-CDNA
synthesis kit (Stratagene) and superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 219 a 238 c 302 g 207 t
ORIGIN
Query Match 33.8%; Score 826; DB 13; Length 966;
Best Local Similarity 95.9%; Pred. No. 2.1e-209;

Matches 913; Conservative 0; Mismatches 30; Indels 9; Gaps 6;			
QY 1440	TAAAGCAGTGGCAGCTCCCACTCCGCGCAGGCGAGAGTGTCTCTCATTTATGTG	1499	
Db 952	TAAAGCAGGCGAGCTCCAC-GCGCGCAGGCGAGCGAGTGTCTCTCATAGGGA	894	
QY 1500	ATTCTCTGGAGCCTGTGCCTATGTCACTGC-AATTTTCAGGAGACATATTCAACTCT-CT	1557	
Db 893	TTCTCTGGAAGCCTGTGCCTATGTCACTGCAGGCAATCTTCAGGAGACATAGTCACTCTCT	834	
QY 1558	GCTCTTCCAAA-----CTGGTGTCTATCCGGC--AGAGGGAGGAAGCAGCAAGTGGAG	1612	
Db 833	GCTCTTCCCAAGACCTGGGTGTCTATCCGCCAGAGAGGAGGAGCAAGTGGAG	774	
QY 1613	CAGGCTTACGAGGACGCTTCTGGCCAGAGCACCACCTCTGCTCCAGCGGGCAGCC	1672	
Db 773	CAGGCTTACGAGGACGCTTCTGGCCAGAGCACCACCTCTGCTCCAGCGGGCAGCC	714	
QY 1673	CCTCATAGGAGGCCAGGAGGCCCAAGGCGTAGAGCCTTGTGTAAGCTGTGAGCA	1732	
Db 713	CCTCATAGGAGGCCAGGAGGCCCAAGGCGTAGAGCCTTGTGTAAGCTGTGAGCA	654	
QY 1733	GGAGAGCGGTGCCACACAGCTTCAGGTCTCCGTGACCTGCTCTATGCCCCACAC	1792	
Db 653	GGAGAGCGGTGCCACACAGCTTCAGGTCTCCGTGACCTGCTCTATGCCCCACAC	594	
QY 1793	CCTACGTGGCGTGGCTCTGTGAGTGTACGTAGTACTCGCTGCTGGTGTCTGTCT	1852	
Db 593	CCTACGTGGCGTGGCTCTGTGAGTGTACGTAGTACTCGCTGCTGGTGTCTGTCT	534	
QY 1853	TGTCGTGAAAGCTTAATGGGTGGCAGGCTGTGTCACTTCTCCAAGCAAGCCATAT	1912	
Db 533	TGTCGTGAAAGCTTAATGGGTGGCAGGCTGTGTCACTTCTCCAAGCAAGCCATAT	474	
QY 1913	GGAGACTATACCACTCCCACTCTGCACACACTCACTCCACCTCTCAAGCCTCCAC	1972	
Db 473	GGAGACTATACCACTCCCACTCTGCACACACTCACTCCACCTCTCAAGCCTCCAC	414	
QY 1973	CTCTGGCCAGATTGGGCTCAATTAATGCTGTGGCTGGCCCTCATCATGAATGACAGCA	2032	
Db 413	CTCTGGCCAGATTGGGCTCAATTAATGCTGTGGCTGGCCCTCATCATGAATGACAGCA	354	
QY 2033	GCTCCCATGGTGGTCTGTGAGCTCTTCAAGTCTTAAATCCCTTAACCTCCAGGATTAG	2092	
Db 353	GCTCCCATGGTGGTCTGTGAGCTCTTCAAGTCTTAAATCCCTTAACCTCCAGGATTAG	294	
QY 2093	CTCCCAAGTGGCTGAGACCCAGCCAGCACACTTCTGGCCCTTCTCCCTGCCTCAATCTA	2152	
Db 293	CTCCCAAGTGGCTGAGACCCAGCCAGCACACTTCTGGCCCTTCT--CTGCCTCAATCTA	235	
QY 2153	AAAGCAGTCCACACCCCTCCAAGTGGATAGAAAGATTTCATGAGTAAGGCTGCAG	2212	
Db 234	AAAGCAGTCCACACCCCTCCAAGTGGATAGAAAGATTTCATGAGTAAGGCTGCAG	175	
QY 2213	GAATCTTATCTCTGGCCACATGTCTCTCCGTGCACACACCAATGAGATTAACTTTGGAAG	2272	
Db 174	GAATCTTATCTCTGGCCACATGTCTCTCCGTGCACACACCAATGAGATTAACTTTGGAAG	115	
QY 2273	TTGACTATTTTATGTCTGCCAGGAGTTCTAATCTGCTCTGTTCCTCTCTCTCTCT	2332	
Db 114	TTGACTATTTTATGTCTGCCAGGAGTTCTAATCTGCTCTGTTCCTCTCTCTCTCTCT	55	
QY 2333	GAAGTCCAGCACACCATTTCTCTCTCCAGTTTCTCTGCCCTCCACCC	2384	
Db 54	GAAGTCCAGCACACCATTTCTCTCTCCAGTTTCTCTGCCCTCCACCC	3	

RESULT 3

BI084101 1048 bp mRNA linear EST 20-JUN-2001
LOCUS 602869466F1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013856 5',
DEFINITION mRNA sequence.
ACCESSION BI084101

VERSION	BI084101.1	GI:14502431
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 1048)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: ccapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: CLONETECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: L1CM1820 row: d column: 17 High quality sequence stop: 853.	
FEATURES	Location/Qualifiers	
source	1..1048	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:5013856"	
	/clone_lib="NIH_MGC_102"	
	/tissue_type="epidermoid carcinoma, cell line"	
	/lab_host="DH10B (phage-resistant)"	
	/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
BASE COUNT	231 a	298 c 264 g 254 t 1 others
ORIGIN	Query Match 32.8%; Score 801.4; DB 13; Length 1048; Best Local Similarity 93.1%; Pred. No. 8.1e-203; Matches 885; Conservative 0; Mismatches 56; Indels 10; Gaps 4;	
QY 1425	GGTATGTTACCTAAAGCCAGTGGCAGCTCCCACTCCGCGCAGGCGAGACTGGAGT	1484
Db 2	GGTATGTTACCTAAAGCCAGTGGCAGCTCCCACTCCGCGCAGGCGAGACTGGAGT	61
QY 1485	CTGTCTCATTTATGTATTCCTGGAGCCTGTGCCTATGTCACTGCAATTTTCAGGAGACAT	1544
Db 62	CTGTCTCATTTATGTATTCCTGGAGCCTGTGCCTATGTCACTGCAATTTTCAGGAGACAT	121
QY 1545	ATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCCGCGAGAGGAGGAGGAGGAGC	1604
Db 122	ATTCAACTCTCTCTCTTCCAAACCTGGTGTCTATCCGCGAGAGGAGGAGGAGGAGC	181
QY 1605	AAGTGGAGCAGGCTTAGCAGGACAGTTCTGCGCAGAACACACAGCTCTGCCAGCG	1664
Db 182	AAGTGGAGCAGGCTTAGCAGGACAGTTCTGCGCAGAACACACAGCTCTGCCAGCG	241
QY 1665	GGGAGCCCCCATAGGAGGCCAGAGGAGGCCCAAGGCGTAGAAGCCTTGTGGAAGC	1724
Db 242	GGGAGCCCCCATAGGAGGCCAGAGGAGGCCCAAGGCGTAGAAGCCTTGTGGAAGC	301
QY 1725	TGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCTTCCCTGACCTGCCCTGTATG	1784
Db 302	TGTGAGCAGGAGAGCGGTGCCACACAGCTTCCAGGTCTTCCCTGACCTGCCCTGTATG	361
QY 1785	CCCAACACCCCTACGTGCCGTGGCTCTGTGAGTGTACGTAGATAGCTCTGCCCTGGGTCT	1844
Db 362	CCCAACACCCCTACGTGCCGTGGCTCTGTGAGTGTACGTAGATAGCTCTGCCCTGGGTCT	421
QY 1845	GTGCTGTTTGTCTGTAAGAGCTTAATGGGCTGGCCAGGCTGTCTCACCTTCTCCAGCAA	1904


```
LOCUS      BI824483              740 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603038855F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179957 5',
mRNA sequence.
ACCESSION  BI824483
VERSION    BI824483.1  GI:15936033
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            NIH-MGC http://mgc.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11448 row: m column: 14
            High quality sequence stop: 709.
            Location/Qualifiers
            1. .740
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5179957"
            /clone_lib="NIH_MGC_115"
            /lab_host="DH10B"
            /note="Organ: pooled brain, lung, testis; Vector:
            pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
            source anonymous pool of 6 male brains, age range 23-27; 1
            male lung, age 27; and 1 male testis, age 69. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.8 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            021. Note: this is a NIH_MGC Library."

BASE COUNT  196 a 181 c 196 g 167 t

ORIGIN
Query Match      30.1%; Score 736.8; DB 13; Length 740;
Best Local Similarity 99.7%; Pred. No. 1.3e-185;
Matches 738; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1  TGGAGTGGGAGCTCAAGCAGGATTCTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 60
Db  1  TGGAGTGGGAGCTCAAGCAGGATTCTCCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT 60
QY  61  CTGAGGCAATGGGTGGAAGGAGGAAGTACTGCAGTTCCTGGAAGAAACAGACCACC 120
Db  61  CTGAGGCAATGGGTGGAAGGAGGAAGTACTGCAGTTCCTGGAAGAAACAGACCACC 120
QY  121  AACATCGGAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCAGAAATT 180
Db  121  AACATCGGAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTTCAGAAATT 180
QY  181  TTCTGTGTGAAGCAAAAGACTGACTGGGAAGCTTTTGTGCTCTGAAGTGCATCAAGAAGTCA 240
Db  181  TTCTGTGTGAAGCAAAAGACTGACTGGGAAGCTTTTGTGCTCTGAAGTGCATCAAGAAGTCA 240
QY  241  CTTGCCCTTCGGGACAGCAGCCCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300
Db  241  CTTGCCCTTCGGGACAGCAGCCCTGGAGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300
QY  301  GAAAACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCTACTACTACCTGGTCATG 360
Db  301  GAAAACATTGTGACCTGGAGGACATCTATGAGAGCACCACCCTACTACTACCTGGTCATG 360

LOCUS      BI818261              742 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603032510F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173587 5',
mRNA sequence.
ACCESSION  BI818261
VERSION    BI818261.1  GI:15928724
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            NIH-MGC http://mgc.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11432 row: d column: 04
            High quality sequence stop: 742.
            Location/Qualifiers
            1. .742
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5173587"
            /clone_lib="NIH_MGC_115"
            /lab_host="DH10B"
            /note="Organ: pooled brain, lung, testis; Vector:
            pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
            source anonymous pool of 6 male brains, age range 23-27; 1
            male lung, age 27; and 1 male testis, age 69. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.8 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            021. Note: this is a NIH_MGC Library."

FEATURES             source
LOCUS      BI818261              742 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603032510F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173587 5',
mRNA sequence.
ACCESSION  BI818261
VERSION    BI818261.1  GI:15928724
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            NIH-MGC http://mgc.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11432 row: d column: 04
            High quality sequence stop: 742.
            Location/Qualifiers
            1. .742
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5173587"
            /clone_lib="NIH_MGC_115"
            /lab_host="DH10B"
            /note="Organ: pooled brain, lung, testis; Vector:
            pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
            source anonymous pool of 6 male brains, age range 23-27; 1
            male lung, age 27; and 1 male testis, age 69. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.8 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            021. Note: this is a NIH_MGC Library."
QY  361  CAGCTTCTTCTGTTGGGGAGCTCTTTTGACCGGATCCTGAGCGGGGTGTTACACAGAG 420
Db  361  CAGCTTCTTCTGTTGGGGAGCTCTTTTGACCGGATCCTGAGCGGGGTGTTACACAGAG 420
QY  421  AAGGATCCAGCTCTGTTGATCCAGCAGGTCTTGTTCGCGACGTGAATATACCTACATGAGAT 480
Db  421  AAGGATCCAGCTCTGTTGATCCAGCAGGTCTTGTTCGCGACGTGAATATACCTACATGAGAT 480
QY  481  GGCATGTCACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCTTACCTCGAAGAAC 540
Db  481  GGCATGTCACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCTTACCTCGAAGAAC 540
QY  541  TCTAAGATCATGATCACTGACTTTGGTCTGTCCCAAGATGGAACAGATGTCATCATGTC 600
Db  541  TCTAAGATCATGATCACTGACTTTGGTCTGTCCCAAGATGGAACAGATGTCATCATGTC 600
QY  601  ACTGCCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCTCAGAC 660
Db  601  ACTGCCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCTCAGAC 660
QY  661  AAGGCTGTGGATTGCTGTGTCATCGGCTCATCACCTACATATTTGCTTGTGGATACCC 720
Db  661  AAGGCTGTGGATTGCTGTGTCATCGGCTCATCACCTACATATTTGCTTGTGGATACCC 720
QY  721  CCGTCTCTATGAAGAACGGA 740
Db  721  CCGTCTCTATGAAGAACGGA 740

RESULT 6
BI818261
LOCUS      BI818261              742 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603032510F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173587 5',
mRNA sequence.
ACCESSION  BI818261
VERSION    BI818261.1  GI:15928724
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            NIH-MGC http://mgc.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11432 row: d column: 04
            High quality sequence stop: 742.
            Location/Qualifiers
            1. .742
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5173587"
            /clone_lib="NIH_MGC_115"
            /lab_host="DH10B"
            /note="Organ: pooled brain, lung, testis; Vector:
            pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
            source anonymous pool of 6 male brains, age range 23-27; 1
            male lung, age 27; and 1 male testis, age 69. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.8 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            021. Note: this is a NIH_MGC Library."
QY  361  CAGCTTCTTCTGTTGGGGAGCTCTTTTGACCGGATCCTGAGCGGGGTGTTACACAGAG 420
Db  361  CAGCTTCTTCTGTTGGGGAGCTCTTTTGACCGGATCCTGAGCGGGGTGTTACACAGAG 420
QY  421  AAGGATCCAGCTCTGTTGATCCAGCAGGTCTTGTTCGCGACGTGAATATACCTACATGAGAT 480
Db  421  AAGGATCCAGCTCTGTTGATCCAGCAGGTCTTGTTCGCGACGTGAATATACCTACATGAGAT 480
QY  481  GGCATGTCACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCTTACCTCGAAGAAC 540
Db  481  GGCATGTCACAGAGACTTAAAGCCCGAAACCTGCTTACCTTACCTTACCTCGAAGAAC 540
QY  541  TCTAAGATCATGATCACTGACTTTGGTCTGTCCCAAGATGGAACAGATGTCATCATGTC 600
Db  541  TCTAAGATCATGATCACTGACTTTGGTCTGTCCCAAGATGGAACAGATGTCATCATGTC 600
QY  601  ACTGCCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCTCAGAC 660
Db  601  ACTGCCTGTGGGACCCAGGCTACGTGGCTCCAGAAAGTGTGGCCAGAAACCTCAGAC 660
QY  661  AAGGCTGTGGATTGCTGTGTCATCGGCTCATCACCTACATATTTGCTTGTGGATACCC 720
Db  661  AAGGCTGTGGATTGCTGTGTCATCGGCTCATCACCTACATATTTGCTTGTGGATACCC 720
QY  721  CCGTCTCTATGAAGAACGGA 740
Db  721  CCGTCTCTATGAAGAACGGA 740
```

BASE COUNT ORIGIN	197 a	180 c	197 g	168 t
Query Match	29.5%; Score 720.8; DB 13; Length 742;			
Best Local Similarity	99.6%; Pred. No. 2.4e-181;			
Matches 733; Conservative	0; Mismatches 2; Indels 1; Gaps 1;			
Qy	1	TGGAGTGGAGGCTCAA- 	GCAGGATTCTTCCGAGTTCCTCGAGTCCCTGCATCCCTCAGAAGCTTCAAC	59
Db	7	TGGAGTGGAGGCTCAATCAGGATCTTCCGAGTTCCTCGAGTCCCTGCATCCCTCAGAAGCTTCAAC	66	
Qy	60	TCTGGAGCAATGGGTGAAAGGAAAGAGATGACTGCGATTCTTCCTGGGAAGAAACAGACAC	119	
Db	67	TCTGGAGCAATGGGTGAAAGGAAAGAGATGACTGCGATTCTTCCTGGGAAGAAACAGACAC	126	
Qy	120	CAACATCCGGAACCTTTCATTTTATGGAAGTGCTGGAGTCAGGAGCTTCTCAGAAGT	179	
Db	127	CAACATCCGGAACCTTTCATTTTATGGAAGTGCTGGAGTCAGGAGCTTCTCAGAAGT	186	
Qy	180	TTTCTGTGTAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTC	239	
Db	187	TTTCTGTGTAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTC	246	
Qy	240	ACCTGCTTCCGGACACAGCTGGAGAATGAGATTGCTGTGTAAGTGCATCAAGAAGTC	299	
Db	247	ACCTGCTTCCGGACACAGCTGGAGAATGAGATTGCTGTGTAAGTGCATCAAGAAGTC	306	
Qy	300	TGAAACATTTGACCTTGGAGGACATCTATGAGAGCACCAACCTACTACTCTGGTGCAT	359	
Db	307	TGAAACATTTGACCTTGGAGGACATCTATGAGAGCACCAACCTACTACTCTGGTGCAT	366	
Qy	360	GCAGCTTTGCTGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGGTGTCTACACAGA	419	
Db	367	GCAGCTTTGCTGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGGTGTCTACACAGA	426	
Qy	420	GAAGATCCAGTCTGGTGATCCAGCAGTCTTCTGCGAGTGAATACCTACATGAGAA	479	
Db	427	GAAGATCCAGTCTGGTGATCCAGCAGTCTTCTGCGAGTGAATACCTACATGAGAA	486	
Qy	480	TGGCATCTGTCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGGAA	539	
Db	487	TGGCATCTGTCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGGAA	546	
Qy	540	CTCTAAGATCATGATGACTTGGTCTGTCCAGATGGAACAGATGCGATCATGTC	599	
Db	547	CTCTAAGATCATGATGACTTGGTCTGTCCAGATGGAACAGATGCGATCATGTC	606	
Qy	600	CACTGCTGTGGACCCAGGCTACGTGCTCCAGAGTGTGCGCCAGAAACCTACAG	659	
Db	607	CACTGCTGTGGACCCAGGCTACGTGCTCCAGAGTGTGCGCCAGAAACCTACAG	666	
Qy	660	CAAGGCTGTGGATTGCTGGTCCATCGGCTCATCACCTACATATTTGCTCTGGATACCC	719	
Db	667	CAAGGCTGTGGATTGCTGGTCCATCGGCTCATCACCTACATATTTGCTCTGGATACCC	726	
Qy	720	CCGTTCTATGAAGAA	735	
Db	727	CCGTTCTATGAAGAA	742	
RESULT 7				
BM921532				
LOCUS	BM921532 1068 bp mRNA linear EST 12-MAR-2002			
DEFINITION	AGENCOURT_6708041 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753010			
	5', mRNA sequence.			
ACCESSION	BM921532			
VERSION	BM921532.1 GI:19371911			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1068)			

AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12788 row: b column: 19
High quality sequence stop: 592.
Location/Qualifiers

FEATURES

source

1. .1068
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5753010"
/lab_host="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT
ORIGIN

Query Match 29.4%; Score 720.4; DB 14; Length 1068;
Best Local Similarity 98.9%; Pred. No. 3.8e-181;
Matches 746; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Qy	1	TGGAGTGGGAGCTCAAGCAGGATTCTTCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT	60
Db	14	TGGAGTGGGAGCTCAAGCAGGATTCTTCCGAGTCCCTGGCATCCTCAGAAGCTTCAACT	73
Qy	61	CTGAGGCAATGGTCCAAAGGAAGAGATGACTGCAGTTCTCTGGGAAGAACAGACCACC	120
Db	74	CTGAGGCAATGGTCCAAAGGAAGAGATGACTGCAGTTCTCTGGGAAGAACAGACCACC	133
Qy	121	AACATCCGGAACCTTTCATTTTATGGAAGTGTCTGGGATCAGGAGCTTTCTCAGAAGTT	180
Db	134	AACATCCGGAACCTTTCATTTTATGGAAGTGTCTGGGATCAGGAGCTTTCTCAGAAGTT	193
Qy	181	TTCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAGAAGTCA	240
Db	194	TTCTGTGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAGAAGTCA	253
Qy	241	CCTGCCCTCCGGACACAGCAGCTGGGAATGAGATTGCTGTGTAAGAAAGATCAAGCAT	300
Db	254	CCTGCCCTCCGGACACAGCAGCTGGGAATGAGATTGCTGTGTAAGAAAGATCAAGCAT	313
Qy	301	GAAGCAATTGTGACCTTGGAGGACATCTATGAGAGCACCAACCCACTACTACCTGGTCATG	360
Db	314	GAAGCAATTGTGACCTTGGAGGACATCTATGAGAGCACCAACCCACTACTACCTGGTCATG	373
Qy	361	CAGCTTGTCTTGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCTTACACAGAG	420
Db	374	CAGCTTGTCTTGTGGGAGCTCTTTGACCGGATCCTGGAGCGGGGTGCTTACACAGAG	433
Qy	421	AAGATGCCAGTCTGGTGATCCAGCAGCTTGTGCGCAGTGAATACCTACATGAGAAT	480
Db	434	AAGATGCCAGTCTGGTGATCCAGCAGCTTGTGCGCAGTGAATACCTACATGAGAAT	493
Qy	481	GGCATCTGCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC	540
Db	494	GGCATCTGCACAGAGACTTAAAGCCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC	553

```
QY 541 TCTAAGATCATGATGACTGTTGGTCTGTCCAGATGAACAGAAATGGCATCATGTCC 600
Db 554 TCTAAGATCATGATGACTGTTGGTCTGTCCAGATGAACAGAAATGGCATCATGTCC 613
QY 601 A-CTGGCTGTGGACCCAGGCTACCTGGCTCCAGAAAGTGTGGCCAGAAACCCCTACAG 659
Db 614 ACCTGCTGTGGACCCAGGCTACCTGGCTCCAGAAAGTGTGGCCAGAAACCCCTACAG 673
QY 660 CAAGGCTGTGGATGTGTGTCTCCATCGGCTCATCACTACATATATGTCTGTGGATACCC 719
Db 674 CAAGGCTGTGGATGTGTGTCTCCATCGGCTCATCACTACATATATGTCTGTGGATACCC 733
QY 720 CCGG-TTCTATGAAGAAACGGAGTCTTAAGCTTTT 752
Db 734 CCCGCTCCTATGAAGAAACGGATCTTAACCTTTT 767

RESULT 8
BI772626 812 bp mRNA linear EST 25-SEP-2001
LOCUS 603060879F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5210299 5',
DEFINITION mRNA sequence.
ACCESSION BI772626
VERSION BI772626
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Plrimates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 812)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11527 row: m column: 20
High quality sequence start: 2
High quality sequence stop: 808.
Location/Qualifiers
1. 812
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5210299"
/lab_host="NIH_MGC_122"
/lab_host="PH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site.1: NotI; Site.2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 217 a 186 c 221 g 188 t
ORIGIN

Query Match 29.1%; Score 713.2; DB 13; Length 812;
Best Local Similarity 91.0%; Pred. No. 2.7e-179;
Matches 801; Conservative 0; Mismatches 3; Indels 76; Gaps 1;

QY 1 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTCAGAACCTTCAACT 60
Db 9 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTCAGAACCTTCAACT 68
```

```
QY 61 CTGGAGGCAATGGGTGGAAGAAAGATGACTGCTAGTTCTCTGGAAGAAACAGACACC 120
Db 69 CTGGAGGCAATGGGTGGAAGAAAGATGACTGCTAGTTCTCTGGAAGAAACAGACACC 128
QY 121 AACATCCGGAACAACTTCATTTTATGGAAGTCTGGATCAGGAGCTTTCTCAGAAGTT 180
Db 129 AACATCCGGAACAACTTCATTTTATGGAAGTCTGGATCAGGAGCTTTCTCAGAAGTT 188
QY 181 TTCTGTGTGAAGAAAGAACTGACTGGGAAGCTCTTTCTCTGAAGTGCATCAAGAACTCA 240
Db 189 TTCTGTGTGAAGAAAGAACTGACTGGGAAGCTCTTTCTCTGAAGTGCATCAAGAACTCA 248
QY 241 CCTGCTTCCGGGACAGCAGCTGGAGATGAGATTGCTGTGTTGAAAAGATCAAGCAT 300
Db 249 CCTGCTTCCGGGACAGCAGCTGGGAATGAGATTGCTGTGTTGAAAAGATCAAGCAT 308
QY 301 GAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCACCTACTACTACCTGGTCATG 360
Db 309 GAAACATTTGTGACCTGGAGGACATCTATGAGAGCACCACCTACTACTACCTGGTCATG 368
QY 361 CAGCTTGTCTGTGTGGGAGCTCTTTGACCGGATCTCGGAGCGGGGTGTCTACACAGAG 420
Db 369 CAGCTTGTCTGTGTGGGAGCTCTTTGACCGGATCTCGGAGCGGGGTGTCTACACAGAG 428
QY 421 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTGTGCGCAGTGAATACCTACATGAAAT 480
Db 429 AAGGATGCCAGTCTGGTGATCCAGCAGGTCTTGTGCGCAGTGAATACCTACATGAAAT 488
QY 481 GGCATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC 540
Db 489 GGCATCGTCCACAGAGACTTAAAGCCGAAACCTGCTTTACCTTACCCCTGAAGAGAAC 548
QY 541 TCTAAGATCATGATGACTGTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCC 600
Db 549 TCTAAGATCATGATGACTGTTGGTCTGTCCAAAGATGGAACAGATGGCATCATGTCC 608
QY 601 ACTGCTGTGGAGCCCGAGCTAGTGTGCTCCAGAGTGTGGCCAGAAACCCCTACAGC 660
Db 609 ACTGCTGTGGAGCCCGAGCTAGTGTGCTCCAGAGTGTGGCCAGAAACCCCTACAGC 636
QY 661 AAGGCTGTGGATTTGCTGTGCTCCATCGGCTCATCATCATATATTGCTCTGTGGATACCC 720
Db 637 -----GCTCTGTGGATACCC 652
QY 721 CCGTCTTATGAAGAAACGGAGTCTAAGCTTTTCGAGAGATCAAGAGGGCTACTATGAG 780
Db 653 CCATTCCTATGAAGAAACGGAGTCTAAGCTTTTCGAGAGATCAAGAGGGCTACTATGAG 712
QY 781 TTTGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGCAAGGACTTTATTGTCACCTTG 840
Db 713 TATGAGTCTCCATTTCTGGGATGACATTTCTGAGTCAGCAAGGACTTTATTGTCACCTTG 772
QY 841 CTTGAGAAGGATCCGACAGCAGCGGTACACCTGTGAGAAGG 880
Db 773 CTTGAGAAGGATCCGACAGCAGCGGTACACCTGTGAGAAGG 812

RESULT 9
BI772626 1129 bp mRNA linear EST 05-MAR-2002
LOCUS 6575096 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5732298
DEFINITION 5', mRNA sequence.
ACCESSION BI772626
VERSION BI772626
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1129)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12734 row: c column: 19
High quality sequence start: 11
High quality sequence stop: 651.
Location/Qualifiers
1. .1129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5732298"
/tissue_lib="NIH_MGC_124"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcorV
(destroyed); Site_2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcorV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
BASE COUNT 287 a 308 c 294 g 240 t

BASE COUNT
ORIGIN

Query Match 29.0%; Score 710.2; DB 14; Length 1129;
Best Local Similarity 95.1%; Pred. No. 2.1e-178;
Matches 800; Conservative 0; Mismatches 28; Indels 13; Gaps 6;

QY 1 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTCAGAACTTCAACT 60
|||||
DB 36 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTCAGAACTTCAACT 95
|||||
QY 61 CTGGAGGCAATGGTCCGAAGGAAGATGACTGCAGTCTCTGGAAGAAACAGACCACC 120
|||||
DB 96 CTGGAGGCAATGGTCCGAAGGAAGATGACTGCAGTCTCTGGAAGAAACAGACCACC 155
|||||
QY 121 AACATCCGGAAACCTTCAATTTATGGAAGTCTGGGATCAGGAGCTTCTTCAGAAAGTT 180
|||||
DB 156 AACATCCGGAAACCTTCAATTTATGGAAGTCTGGGATCAGGAGCTTCTTCAGAAAGTT 215
|||||
QY 181 TTCTCTGGTGAAGCAAGACAGTCTGGAAGTCTTTGCTCTGAAGTGCATCAAGAAGTCA 240
|||||
DB 216 TTCTCTGGTGAAGCAAGACAGTCTGGAAGTCTTTGCTCTGAAGTGCATCAAGAAGTCA 275
|||||
QY 241 CTGTCCTTCGGGACAGCAGCTGGAGATGAGATGCTGCTTGAAGAAGATCAAGCAT 300
|||||
DB 276 CTGTCCTTCGGGACAGCAGCTGGAGATGAGATGCTGCTTGAAGAAGATCAAGCAT 335
|||||
QY 301 GAAACATTTGTACCTCGGAGGACATCTATGAGAGCACCACCCTACTTACCTGGTCTATG 360
|||||
DB 336 GAAACATTTGTACCTCGGAGGACATCTATGAGAGCACCACCCTACTTACCTGGTCTATG 395
|||||
QY 361 CAGCTTTGTTGGTGGGAGCTCTTTGACCGGATCTCTGGAGGGGGTGTCTACACAGAG 420
|||||
DB 396 CAGCTTTGTTGGTGGGAGCTCTTTGACCGGATCTCTGGAGGGGGTGTCTACACAGAG 455
|||||
QY 421 AAGGATCCAGTCTGTGTATCCAGCAGCTTGTGCGGAGTCAAAATACCTACATCAGAT 480
|||||
DB 456 AAGGATCCAGTCTGTGTATCCAGCAGCTTGTGCGGAGTCAAAATACCTACATCAGAT 515
|||||
QY 481 GGCATCGTCCACAGAGACTTAAAGCCGAAAAACCTTGTACCTTACCCCTGAAGAGAAC 540
|||||
DB 516 GGCATCGTCCACAGAGACTTAAAGCCGAAAAACCTTGTACCTTACCCCTGAAGAGAAC 575
|||||
QY 541 TCTAAGATCATGATCACTGACTTGGTCTGTCCAGATGGACAGATGGCATCATGTCC 600

Db 576 TCTAAGATCATGATCACTGACTTTGGTCTGTCCAAGATGGAACAGATGCATCATGTCC 635
|||||
QY 601 ACTGCCTGTGGGACCCAGGCTAGCTGGCTCCAGAGTGTGGCCAGAAACCTACAGC 660
|||||
Db 636 ACTGCCTGTGGGACCCAGGCTAGCTGGCTCCAGAGTGTGGCCAGAAACCTACAGC 695
|||||
QY 661 AAGGC---TGTGGATGCTGTGCTCATCGGCGTCAATCA-CCTACATATGCTCTGTGGATA 716
|||||
Db 696 AAGGCCTGTGGATTTGCTGGTCCATCGGGGTCATCACCTTACCTATTTGCTCTGTGGATA 755
|||||
QY 717 ---CCCCCGTCTTATGAA-CAAAACGGAGTCTAAGCTTTTCGAGAGATCAAGGAGGCT 772
|||||
Db 756 ACCCCCCCGTCTTATGAAGGAAACGGAGTCTAAGCTTTTCGAGAGATCAAGGAGGCT 815
|||||
QY 773 ACTATGAG---TTTGAGTCTCCATTTCTGGG---ATGACATTTCTGAGTCAGCCAAAGGACTT 827
|||||
Db 816 TAATATGAGTTTTGAGTCTCCATTCCTGGGATGAACATTCCTGAGTCAGCCCGGAGCT 875
|||||
QY 828 T 828
Db 876 T 876

RESULT 10
Bg715920
LOCUS
DEFINITION 602676667F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799224 5',
mRNA sequence.
ACCESSION Bg715920
VERSION Bg715920.1 GI:13995107
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 824)
AUTHORS NTH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10688 row: e column: 17
High quality sequence stop: 766.
FEATURES
source
1. .824
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4799224"
/tissue_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTN-3',
size-selected for average insert size 2.3 kb and
normalized to 500. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI), National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 223 a 194 c 219 g 188 t
ORIGIN

Query Match

28.7%; Score 701.4; DB 12; Length 824;

Best Local Similarity 96.9%; Pred. No. 4e-176; Matches 747; Conservative 0; Mismatches 21; Indels 3; Gaps 3;

```
QY 1 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTTCAAGAGCTTCAACT 60
Db 38 TGGAGTGGGAGCTCAAGCAGGATCTTCCCGAGTCCCTGGCATCTTCAAGAGCTTCAACT 97
QY 61 CTGAGGCAATGGTCCGAAGGAGAGATGACTGCTGGATCAGGAGTCTTCTGGAAGAACAGACCAC 120
Db 98 CTGAGGCAATGGTCCGAAGGAGAGATGACTGCTGGATCAGGAGTCTTCTGGAAGAACAGACCAC 157
QY 121 AACATCCGAAAAACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAGTTT 180
Db 158 AACATCCGAAAAACCTTCATTTTATGGAAGTCTGGATCAGGAGCTTCTCAGAGTTT 217
QY 181 TTCTGTTGGAAGCAAGACTGCTGGAGAGCTTTTGGTCTCTGAAAGTCAAGAGTCA 240
Db 218 TTCTGTTGGAAGCAAGACTGCTGGAGAGCTTTTGGTCTCTGAAAGTCAAGAGTCA 277
QY 241 CCTGCCCTCCGGGACAGCAGCTGGAGATGAGATGCTGTGTAAGAGATCAAGCAT 300
Db 278 CCTGCCCTCCGGGACAGCAGCTGGAGATGAGATGCTGTGTAAGAGATCAAGCAT 337
QY 301 GAAACATTTGTACCTGGAGGACATCTATGAGAGCACCACCTACTTACCTTGTCTATG 360
Db 338 GAAACATTTGTACCTGGAGGACATCTATGAGAGCACCACCTACTTACCTTGTCTATG 397
QY 361 CAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGTCTTACACAGAG 420
Db 398 CAGCTTGTCTGGTGGGAGCTCTTTGACCGGATCCTGGAGGGGTCTTACACAGAG 457
QY 421 AAGGATGCCAGTCTGGTGTATCCAGCAGGCTTTGTGGCAGTGAATACCTACATGAGAT 480
Db 458 AAGGATGCCAGTCTGGTGTATCCAGCAGGCTTTGTGGCAGTGAATACCTACATGAGAT 517
QY 481 GGCATCGTCCACAGAGACTTAAAGCCCGAAACCTTCTTACCTTACCCCTGAAGAGAAC 540
Db 518 GGCATCGTCCACAGAGACTTAAAGCCCGAAACCTTCTTACCTTACCCCTGAAGAGAAC 577
QY 541 TCTAAGATCATGATCACTGACTTTGGTCTGTCAGAGTGGACAGAAATGCGATCATGTCC 600
Db 578 TCTAAGATCATGATCACTGACTTTGGTCTGTCAGAGTGGACAGAAATGCGATCATGTCC 637
QY 601 ACTGCCCTGGGACCCAGGCTACG-TGGCTCCAGAGTGTGGCCAG-AAACCCCTACA 658
Db 638 ACTGCCCTGGGACCCAGGCTACGTTGGCTCCAGAGTGTGGCCAGAAACCCCTACA 697
QY 659 GCAAGCGTGTGGATGCTGGTCCATCGGCGTCACTACCTTACATATTTGCTCTGTGGATACC 718
Db 698 GCAAGCGTGTGGATGCTGGTCCATCGGCGT-ATCACCTTAACATATGCTCTGTGGATACC 756
QY 719 CCCGCTTCTATGAAGAAACGGAGTCTTAAGCTTTTCGAGAGATCAAGGAGG 769
Db 757 CCCCATCTATGAAGAAACGGAGTCTTACGCTTTTCGAAAAATCAAGAGGG 807
```

RESULT 11

BI753035 731 bp mRNA linear EST 25-SEP-2001
LOCUS 603025844Fl NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196323 5',
DEFINITION mRNA sequence.

ACCESSION BI753035
VERSION BI753035.1 GI:15744613
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 731)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: ccapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11491 row: g column: 12
High quality sequence stop: 729.

FEATURES
source

1..731
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5196323"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

BASE COUNT 190 a 182 c 191 g 168 t
ORIGIN

Query Match 28.6%; Score 699; DB 13; Length 731;
Best Local Similarity 99.7%; Pred. No. 1.6e-175;
Matches 721; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 TGGAGTGGGAGCTCAAG-CAGGATTTCTTCCCGAGTCCCTGGCATCTTCAAGAGCTTCAAC 59
Db 1 TGGAGTGGGAGCTCAAGCCAGGATTTCTCCCGAGTCCCTGGCATCTTCAAGAGCTTCAAC 60

QY 60 TCTGAGGCAATGGTCCGAAGGAGAGATGACTGAGTTCCTGGAGAAACAGACAC 119
Db 61 TCTGAGGCAAT-GGTCGAAGGAGAGATGACTGAGTTCCTGGAGAAACAGACAC 119

QY 120 CAACATCCGAAAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGT 179
Db 120 CAACATCCGAAAAACCTTCATTTTATGGAAGTCTGGGATCAGGAGCTTCTCAGAGT 179

QY 180 TTCTCTGGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTC 239
Db 180 TTCTCTGGTGAAGCAAGACTGACTGGGAAGCTCTTTGCTCTGAAGTGCATCAAGAAGTC 239

QY 240 ACCTGCCCTTCCGGGACAGCAGCTGGGAATGAGATTCGCTGTGGAAGATCAAGCA 299
Db 240 ACCTGCCCTTCCGGGACAGCAGCTGGGAATGAGATTCGCTGTGGAAGATCAAGCA 299

QY 300 TGAAGCAATTCGTGACCTGGAGGACATCTATGAGAGCACCACCCACTACTACCTGGTCA 359
Db 300 TGAAGCAATTCGTGACCTGGAGGACATCTATGAGAGCACCACCCACTACTACCTGGTCA 359

QY 360 GCAGCTTGTCTGGTGGGAGCTTTTGACCGGATCCTGGAGCGGGGTGTCTACAGAA 419
Db 360 GCAGCTTGTCTGGTGGGAGCTTTTGACCGGATCCTGGAGCGGGGTGTCTACAGAA 419

QY 420 GAAGGATGCCAGTCTGGTGTATCCAGAGGCTTCTTCCGCGAGTGAATACCTACATGAGAA 479
Db 420 GAAGGATGCCAGTCTGGTGTATCCAGAGGCTTCTTCCGCGAGTGAATACCTACATGAGAA 479

QY 480 TGGCATCGTCCACAGAGACTTAAAGCCCGAAACCTTGTACCTTACCCCTGAAGAGAA 539
Db 480 TGGCATCGTCCACAGAGACTTAAAGCCCGAAACCTTGTACCTTACCCCTGAAGAGAA 539

QY 540 CTCCTAAGATCATGATCACTGACTTTGGTCTGTCCAAGATGGAACAGATGCAATGTC 599
Db 540 CTCCTAAGATCATGATCACTGACTTTGGTCTGTCCAAGATGGAACAGATGCAATGTC 599

```
QY 600 CACTGCTGTGGGACCCAGGCTAGCTGCTCCAGAACTGCTGGCCAGAAAACCCCTACAG 659
|||||
Db 600 CACTGCTGTGGGACCCAGGCTAGCTGCTCCAGAACTGCTGGCCAGAAAACCCCTACAG 659
QY 660 CAAGCTGTGGATTCTGTGTCATCGGCGTCATCACTACACCTACATATTGCTCTGTGGATACCC 719
|||||
Db 660 CAAGCTGTGGATTCTGTGTCATCGGCGTCATCACTACACCTACATATTGCTCTGTGGATACCC 719
QY 720 CCC 722
|||
Db 720 CCC 722

RESULT 12
BQ934044
LOCUS AGENCOURT_8802332 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6336727
DEFINITION 5', mRNA sequence.
ACCESSION BQ934044
VERSION BQ934044.1 GI:22349427
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Mark Moconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13801 row: d column: 08
High quality sequence stop: 689.
Location/Qualifiers
1..905
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6336727"
/clone_lib="NIH_MGC_130"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Otcysts; Vector: pCMV-SPORT6.1.cdb;
Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally.
Primer: Oligo dt. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is a
NIH_MGC Library."
BASE COUNT 238 a 217 c 240 g 210 t
ORIGIN
Query Match 26.7%; Score 653.6; DB 14; Length 905;
Best Local Similarity 90.0%; Pred. NO. 2.5e-163;
Matches 745; Conservative 0; Mismatches 79; Indels 4; Gaps 4;
QY 31 GAGTCCCTGGCATCTCAGAACTTCACTCTGGAGGCAATGGGTGGAAGGAAGAAGAT 90
|||||
Db 68 GAGCCCTCGCTTCACAGGACCTTCACTCTGGAGGACATGGGCGGAAGAGAGGAG 127
QY 91 GACTGCAGTTCTTGGGAAGAACAGACCCACCAACATCCGGAACACCTTCATTTTATGGAA 150
|||||
Db 128 GACTGCAGTTCTTGGGAAGAACAGACCCACCAACATCAGGAAAACCTTCATCTTCATGGAA 187
QY 151 GTGCTGGGATCAGGAGCTTTCTCAGAAGTTTTCCTGGTGAAGCAAGAGTACTGGGGAAG 210
|||||
Db 188 GTGCTGGGATCAGGAGCTTTCTCAGAAGTTTTCCTGGTGAAGCAAGAGTACTGGGGAAG 247
QY 211 CTCTTTGCTCTGAAGTGCATCAAGAAGTACCTGCTTCCTGGGACAGCAGCGCTGGAGAT 270
```

```
Db 248 CTCTTTGCTCTGAATGTATCAAGAAGTCAACAGCCTTCGGGACAGCAGCCCTAGAGAT 307
QY 271 GAGATTTCTGTTGAAAAAGATCAAGCATGAAAAATTTGTGACCCCTGGAGGACATCTAT 330
|||||
Db 308 GAGATCCTCTGTTGAAAAAGATCAAGCATGAAAAATTTGTGACCCCTGGAGGACATCTAT 367
QY 331 GAGAGCACCAACCACTACTACCTGGTGCATGCAGCTGTTTCTGGTGGGAGCTCTTTGAC 390
|||||
Db 368 GAGAGCACCAACCACTACTACCTGGTGCATGCAGCTGTTTCTGGAGTGAGCTTTTGGAC 427
QY 391 CGGATCCTGGAGCGGGGTGTTACACAGAGAAGGATGCCAGTCTGGTGTATCCAGCAGGTC 450
|||||
Db 428 CGGATCCTGGAGCGGTGTTGTTACACAGAAAAGGATGCCAGCTGGTGTATCCAGCAGGTC 487
QY 451 TTGTCGGCAGTGAATAACCTACATGAGAATGGCATCTCCACAGAGACTTTAAAGCCCGAA 510
|||||
Db 488 TTGTCGGCAGTGAATAACCTTCATGAGAATGGCATCTCCACAGAGATCTAAAGCCCTGAA 547
QY 511 AACCTGCTTACCTTACCCCTGAAGAACTCTAAGATCATGATCACTGACTTTGGTCTG 570
|||||
Db 548 AACCTGCTTACCTTACCCCTGAGGAGAACTCCAAGATCATGATCACTGACTTTGGTCTA 607
QY 571 TCCAGATGGAACAGATGCATCATCTCCACTGCCCTGCTGGACCCCGAGCTAGTGGCT 630
|||||
Db 608 TCCAGATGGAAGAGAAATGGAGTATGTCCACAGCTTGTGGACCCCGAGCTAGTGGCT 667
QY 631 CCAGAAGTGTGTGGCCCAAGAAACCTCACAGCAAGCTGTGGATTGCTGGTCCATCGCGCTC 690
|||||
Db 668 CCAGAAGTGTGTGGCCCAAGAAACCTCACAGTAAGCTGTGGACTGCTGCTCATTTGGTGC 727
QY 691 ATCACTACATATTTGCTCTGTGGATACCCCGCTTCTATGAAGAACGGAGTCTAAGC-T 749
|||||
Db 728 ATCACTACATATTTGCTGTGGCTATCCCTTCTATGAAGAAACAGAAATCAAGCTT 787
QY 750 TTTCGAGAAGATCAAGGAGGCTACTATGA-GTTTGTAGTCTCCATTTCTGGGATGACATTT 808
|||||
Db 788 TTTTGAAGAATCAAAAGAGGTTACTACAGAGTTTGTAGTCTCCGTTCTGGATGACATTT 847
QY 809 CTGAGTCA-GCCAAAGGACTTTATTGCCA-CTTGTCTTGAGAAGATCC 854
|||||
Db 848 CTGAGTCAGCCAGGGATTTTATTGTCATCTCTGCTGGAGAAGGACCC 895

RESULT 13
BQ293660
LOCUS 602390529F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502479 5',
DEFINITION mRNA sequence.
ACCESSION BQ293660
VERSION BQ293660.1 GI:13053536
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10371 row: i column: 08
High quality sequence stop: 765.
Location/Qualifiers
1..818
FEATURES
source
```

REFERENCE
AUTHORS
1 (bases 1 to 650)
NIH-MGC <http://mgc.nci.nih.gov/>.

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cygaps-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshuyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://Image.llnl.gov>
 Plate: LHAM1791 row: i column: 19
 High quality sequence stop: 643.
 Location/Qualifiers
 1..656
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5312346"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /notes="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-xhoI (gtcgaag
 p); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTNN-3',
 size-selected for average insert size 2.3 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 177 a 156 c 177 g 146 t
 ORIGIN
 Query Match 22.7% Score 555: DB 13: Length 656;
 Best Local Similarity 97.6% Pred. No. 5e-137;
 Matches 606; Conservative 0; Mismatches 10; Indels 5; Gaps 4;
 QY 1 TGGAGTGGGAGCTCAA-GCAGGATCTTCCCGAGTCCCTGGCATCCTCAGAAAGCTTCAAC 59
 |||||
 Db 36 TGGAGTGGGAGCTCAACGACGAGATCTTCCGAGTCCCTGGCATCCTCAGAAAGCTTCAAC 95
 |||||
 QY 60 TCTGGAGGCAATGGTGCAGGAGGAAGATGACTGCGAGTTCCTGGAGAAACAGACAC 119
 |||||
 Db 96 TCTGGAGGCAATGGTGCAGGAGGAAGATGACTGCGAGTTCCTGGAGAAACAGACAC 155
 |||||
 QY 120 CAACATCGGAAAAACCTTCATTTTATGGAAGTCTGGGATCAGAGGCTTCTCAGAGT 179
 |||||
 Db 156 CAACATCGGAAAAACCTTCATTTTATGGAAGTCTGGGATCAGAGGCTTCTCAGAGT 215
 |||||
 QY 180 TTTCTGTTGAAGCAAAAGACTGACTGGGAAGCTTTTGCTCTGAAGTGCATCAAGAAGTC 239
 |||||
 Db 216 TTTCTGTTGAAGCAAAAGACTGACTGGGAAGCTTTTGCTCTGAAGTGCATCAAGAAGTC 275
 |||||
 QY 240 ACCTGCCTTCGGGACAGACACCGCTGGGAATGAGATTGCTGTGTGAAAAAGATCAAGCA 299
 |||||
 Db 276 ACCTGCCTTCGGGACAGACACCGCTGGGAATGAGATTGCTGTGTGAAAAAGATCAAGCA 335
 |||||
 QY 300 TGAACAATGTGACCCCTGGAGGACATCTATGAGAGCACCAACCCACACTACTACCTGGTCA 359
 |||||
 Db 336 TGAACAATGTGACCCCTGGAGGACATCTATGAGAGCACCAACCCACACTACTACCTGGTCA 395
 |||||


```
QY 360 GCAGCTTCTTCTGGTGGGAGCTCTTTGACCGATCC--TGGAGCGGGTGTCTACACA 417
Db 396 GCAGCTTCTTCTGGTGGGAGCTCTTTGACCGGATCATGGCAGCGGGTGTCTACACA 455
QY 418 GAGAGGATGCCAGTCTGGTGATCCAGCAGGTCTTGTGGCAGTGAATACCTACATGAG 477
Db 456 GAGAGGATGCCAGTCTGGTGATCCAGCAGGTCTTGTGGCAGTGAATACCTACATGAG 515
QY 478 AATGGCATGTCACAGAGACTTAAAGCCCG-AAAACCTGCTTTACTTTACCCCTGAAGA 536
Db 516 AATGGCATGTCACAGAGACTTAAAGCCCGAAAAACCTGCTTTACTTTACCCCTGAAGA 575
QY 537 GAACCTTAAGATCATGATCACTGACTTGTGCTGCTCAAGATGGAACAGAAATGGCATC-A 595
Db 576 GAACCTTAGGATCATGATCACTGACTGCTGCTGCTCAAGATGGAACAGAAATGGCATCA 635
QY 596 TGTCCACTGCTGTGGGACCC 616
Db 636 TGTCCACTGCTGTGGGACCC 656

RESULT 15
BI834635 554 bp mRNA linear EST 04-OCT-2001
DEFINITION 603090418F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5229245 5',
mRNA sequence.
ACCESSION BI834635
VERSION BI834635.1 GI:15946185
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 554)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L18M11577 row: c column: 06
High quality sequence stop: 554.
FEATURES
source
1..554
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5229245"
/lab_host="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT 156 a 128 c 143 g 127 t
ORIGIN
Query Match 21.3%; Score 522; DB 13; Length 554;
Best Local Similarity 98.0%; Pred. No. 3.1e-128;
Matches 543; Conservative 0; Mismatches 0; Indels 11; Gaps 1;
QY 82 GAAGAAGATGACTGCTGCTGAGAGAAACAGACCACCAACATCCGGAACCTTCATT 141
```

Search completed: March 14, 2003, 15:22:05
Job time : 3891 secs